PEMODELAN REGRESI LOGISTIK MULTINOMIAL MENGGUNAKAN METODE *PENALIZED MAXIMUM LIKELIHOOD ESTIMATION* UNTUK MENGATASI PEMISAHAN

SKRIPSI

oleh:
ACHMAD FANDY KURNIAWAN
125090501111006

PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS BRAWIJAYA
MALANG
2016

PEMODELAN REGRESI LOGISTIK MULTINOMIAL MENGGUNAKAN METODE *PENALIZED MAXIMUM LIKELIHOOD ESTIMATION* UNTUK MENGATASI PEMISAHAN

SKRIPSI

Sebagai salah satu syarat untuk memperoleh gelar
Sarjana Sains dalam bidang Statistika

oleh:
ACHMAD FANDY KURNIAWAN
125090501111006

PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS BRAWIJAYA
MALANG
2016

LEMBAR PENGESAHAN SKRIPSI

PEMODELAN REGRESI LOGISTIK MULTINOMIAL MENGGUNAKAN METODE *PENALIZED MAXIMUM LIKELIHOOD ESTIMATION* UNTUK MENGATASI PEMISAHAN

oleh:

ACHMAD FANDY KURNIAWAN
125090501111006

Setelah dipertahankan di depan Majelis Penguji
pada tanggal 15 Juni 2016
dan dinyatakan memenuhi syarat untuk memperoleh gelar
Sarjana Sains dalam bidang Statistika

Mengetahui,
Dosen Pembimbing,

Dr. Ir. Maria Bernadetha Mitakda
NIP. 195205211981032001

Ketua Jurusan Matematika
Fakultas MIPA Universitas Brawijaya

Ratno Bagus Edy Wibowo, S.Si, M.Si, Ph.D
NIP. 197509082000031003

LEMBAR PERNYATAAN

Saya yang bertanda tangan di bawah ini :

Nama	**: Achmad Fandy Kurniawan**
NIM	**: 125090501111006**
Jurusan	**: Matematika/Statistika**
Penulis Skripsi Berjudul	**:**

PEMODELAN REGRESI LOGISTIK MULTINOMIAL MENGGUNAKAN METODE *PENALIZED MAXIMUM LIKELIHOOD ESTIMATION* UNTUK MENGATASI PEMISAHAN

Dengan ini menyatakan bahwa :

1. **Isi dari Skripsi ini adalah benar-benar karya saya sendiri dan bukan hasil menjiplak karya orang lain. Karya-karya yang tercantum dalam Daftar Pustaka Skripsi ini semata-mata digunakan sebagai acuan/referensi.**
2. **Apabila di kemudian hari diketahui bahwa isi Skripsi saya merupakan hasil jiplakan, maka saya bersedia menanggung akibat hukum dari keadaan tersebut.**

Demikian pernyataan ini dibuat dengan segala kesadaran.

Malang, 22 Juni 2016
Yang menyatakan,

Achmad Fandy Kurniawan
NIM. 125090501111006

PEMODELAN REGRESI LOGISTIK MULTINOMIAL MENGGUNAKAN METODE *PENALIZED MAXIMUM LIKELIHOOD ESTIMATION* UNTUK MENGATASI PEMISAHAN

ABSTRAK

Parameter model regresi logistik multinomial diduga dengan metode *Maximum Likelihood Estimation* (MLE) dan metode iteratif *Newton-Raphson*. Kadang hasil pendugaan dengan metode MLE bersifat tak hingga atau proses iterasi tidak konvergen dan data mengandung pemisahan sempurna atau kurang sempurna sehingga metode MLE tidak dapat digunakan. Hal ini dapat diselesaikan dengan pendekatan metode *Penalized Maximum Likelihood Estimation* (PMLE) atau Prosedur *Firth* (Firth, 1993). Metode ini menghilangkan bias dengan memodifikasi fungsi skor *likelihood* pada metode MLE menjadi fungsi skor *Penalized Likelihood.* Hasil analisis terhadap data sekunder menunjukkan bahwa data profil pasien poli kardiologi RSUD Saiful Anwar mengandung pemisahan kurang sempurna sehingga digunakan pendekatan metode PMLE. Model terbaik bagi data ini mengandung prediktor usia, tekanan darah *sistole*, tekanan darah *diastole*, kolesterol, LDL dan Trigliserid yang berpengaruh nyata terhadap tingkatan penyakit jantung koroner.

Kata Kunci: PMLE, Pemisahan sempurna atau kurang sempurna, Penyakit Jantung Koroner, Regresi Logistik Multinomial.

MULTINOMIAL LOGISTIC REGRESSION MODELING USING PENALIZED MAXIMUM LIKELIHOOD ESTIMATION METHOD TO SOLVE SEPARATION PROBLEM

ABSTRACT

Paramater of multinomial logistic regression model is estimated using Maximum Likelihood Estimation (MLE) and the Newton-Raphson iterative method. Problems occur when estimators obtained by MLE is infinite or not converge during iteration process and data contains complete separation or quasi-complete separation. This could be solved by Penalized Maximum Likelihood Estimation (PMLE) method or Firth's procedure which was first proposed by Firth (1993). This method will reduce bias by modifying the function score of likelihood in MLE to function score of Penalized Likelihood. The analysis of secondary data showed that patient profile of poly cardiology in Saiful Anwar Hospital contains quasi-complete separation, so PMLE method is used. The best model for the data covered several predictors: age, systolic blood pressure, diastolic blood pressure, cholesterol, LDL and triglycerides were significantly affected levels of coronary heart disease.

Keywords: PMLE, complete separation or quasi-complete separation, coronary heart disease, multinomial logistic regression.

KATA PENGANTAR

Puji syukur ke hadirat Allah SWT atas berkat, rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi berjudul "Pemodelan Regresi Logistik Multinomial Menggunakan Metode *Penalized Maximum Likelihood Estimation* Untuk Mengatasi Pemisahan" sebagai salah satu syarat memperoleh gelar Sarjana Sains dalam bidang Statistika.

Dalam penyusunan skripsi ini, banyak pihak telah memberikan dukungan dan bantuan. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Dr. Ir. Maria Bernadetha Mitakda selaku dosen pembimbing atas waktu, bimbingan, saran dan kesabaran yang telah diberikan kepada penulis selama menyusun skripsi.
2. Rahma Fitriani, S.Si, M.Sc, Ph.D selaku dosen penguji I atas pengarahan dan bimbingan yang telah diberikan.
3. Dr. Umu Sa'adah, M.Si selaku dosen penguji II atas pengarahan dan bimbingan yang telah diberikan.
4. Ratno Bagus E.W, S.Si., M.Sc., Ph.D selaku Ketua Jurusan Matematika Fakultas MIPA Universitas Brawijaya.
5. Ibu, Nanda, Cantika, Emi Setyowati S.Si dan Keluarga Besar di Lampung atas doa dan dukungan yang diberikan kepada penulis dalam menyelesaikan skripsi.
6. Semua pihak yang telah membantu dan memberikan dorongan selama penulisan skripsi ini yang tidak dapat penulis sebutkan satu per satu.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, dengan segala kerendahan hati penulis mengharapkan kritik dan saran yang membangun demi perbaikan skripsi ini. Semoga skripsi ini bermanfaat bagi pembaca.

Malang, 22 Juni 2016

Penulis

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR LAMPIRAN

BAB I
PENDAHULUAN

1.1. Latar Belakang

Regresi merupakan metode yang sangat populer untuk membentuk hubungan fungsi antara dua peubah atau lebih. Analisis ini digunakan untuk menggambarkan fungsi yang menunjukkan arah hubungan antar peubah dan melakukan prediksi. Secara umum, analisis regresi diterapkan untuk menganalisis data dengan peubah respon berupa data kuantitatif. Akan tetapi banyak ditemukan permasalahan di mana peubah respon bersifat kategorik. Permasalahan tersebut dapat diselesaikan dengan menggunakan analisis regresi logistik. Di kehidupan sehari-hari banyak ditemukan masalah bahwa peubah respon pada suatu data kualitatif bersifat nominal, untuk menyelesaikan masalah tersebut, dapat digambarkan ke dalam suatu model yang dikenal dengan model regresi logistik multinomial. Model regresi logistik multinomial merupakan suatu pendekatan model matematis yang dapat digunakan untuk menyatakan hubungan dari beberapa peubah prediktor terhadap peubah respon yang memiliki lebih dari dua kategori (Arieska, 2010).

Parameter model regresi logistik multinomial diduga dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). Kadang hasil pendugaan dengan metode MLE bersifat tak hingga atau proses iterasi tidak konvergen sehingga metode MLE tidak dapat digunakan. Hal tersebut dikarenakan adanya satu atau kombinasi linier beberapa prediktor yang mengakibatkan peubah prediktor dan peubah respon terpisah secara sempurna, serta dapat pula disebabkan oleh ukuran data yang terlalu kecil. Albert and Anderson (1984) menyebut kondisi semacam ini sebagai pemisahan atau *separation*.

Hal ini dapat diselesaikan dengan pendekatan metode *Penalized Maximum Likelihood Estimation* (PMLE) atau Prosedur *Firth* (Firth, 1993). Metode ini menghilangkan bias dengan memodifikasi fungsi skor *likelihood* pada metode MLE menjadi fungsi skor *Penalized Likelihood.*

Metode tersebut akan diterapkan pada data profil pasien kardiologi RSSA Malang. Pada data tersebut, selain pasien yang menderita penyakit jantung koroner (PJK) terdapat juga pasien yang menderita PJK disertai hipertensi dan pasien yang menderita PJK disertai hipertensi dan diabetes mellitus. Oleh karena itu, melalui tugas akhir ini penulis ingin menerapkan metode PMLE pada data profil pasien kardiologi berdasarkan faktor usia, tekanan darah dan profil lemak darah.

1.2. Rumusan Masalah

Rumusan masalah penelitian ini adalah:

1. Bagaimana model regresi logistik multinomial pada data profil pasien kardiologi RSSA Malang ?
2. Faktor-faktor apa yang berpengaruh secara nyata terhadap tingkatan penyakit jantung koroner ?
3. Bagaimana penerapan metode PMLE pada model regresi logistik multinomial untuk mengatasi pemisahan?

1.3. Tujuan Penelitian

Tujuan penelitian ini adalah:

1. Membentuk model regresi logistik multinomial pada data profil pasien kardiologi RSSA Malang.
2. Mengetahui faktor-faktor yang berpengaruh secara nyata terhadap tingkatan penyakit jantung koroner.
3. Menerapkan metode PMLE pada model regresi logistik multinomial untuk mengatasi pemisahan.

1.4. Manfaat Penelitian

Manfaat penelitian ini adalah:

1. Mengetahui bahwa masalah pemisahan dalam analisis regresi logistik multinomial dapat diatasi menggunakan metode PMLE.
2. Mengetahui faktor-faktor yang berpengaruh terhadap tingkatan penyakit jantung koroner.

1.5. Batasan Masalah

Batasan masalah penelitian ini adalah:

1. Peubah respon memiliki tiga kategori.
2. Pendugaan parameter regresi logistik menggunakan metode PMLE untuk mengatasi masalah pemisahan.
3. Data pada penelitian ini dianggap memenuhi asumsi non-multikolinieritas.

BAB II
TINJAUAN PUSTAKA

2.1. Sebaran Multinomial

Sebaran multinomial sering digunakan dalam analisis data kategori. Peubah respon yang memiliki lebih dari dua kategori akan mengikuti sebaran multinomial. Misal terdapat *q* kategori pada peubah respon dengan peluang $\{\pi_1, \pi_2, \ldots, \pi_q\}$ dan $\sum_{k=1}^{q} \pi_k = 1$. Untuk *n* contoh, peluang multinomial bahwa n_1 termasuk kategori 1, n_2 termasuk kategori 2, ..., n_q termasuk kategori *q* di mana $\sum_{k=1}^{q} n_k = n$:

$$P(n_1, n_2, \ldots, n_q) = \left(\frac{n!}{n_1! n_2! \ldots n_q!}\right) {\pi_1}^{n_1} {\pi_2}^{n_2} \ldots {\pi_q}^{n_q} \tag{2.1}$$

untuk kategori ke-*k* di mana *k* = *1,2, ...,q* (Agresti, 2002).

2.2. Regresi Logistik Multinomial

Regresi logistik multinomial diterapkan pada peubah respon yang memiliki lebih dari dua kategori berskala nominal. Apabila terdapat *q* kategori pada peubah respon maka akan terbentuk (*q-1*) fungsi logit. Hosmer dan Lemeshow (2000) menyatakan bahwa untuk model regresi logistik multinomial dengan peubah respon tiga kategori yang diberi kode 1, 2 dan 3 akan didapatkan dua fungsi logit. Sebelum itu, ditentukan terlebih dahulu kategori peubah respon mana yang digunakan sebagai pembanding, kategori respon mana pun dapat menjadi pembanding. Pemilihan kategori pembanding bergantung pada tujuan yang ingin dicapai. Namun, pada umumnya digunakan Y = 1 sebagai pembanding dikarenakan model regresi logistik multinomial merupakan perluasan dari model regresi logistik biner.

Pandang tiga kategori respon *Y* (*k* = *1,2,3*) di mana *Y* = 1 sebagai pembanding, maka model umum regresi logistik multinomial untuk *p* peubah prediktor yang dinyatakan dalam vektor $\boldsymbol{x}$ serta peluang kategori respon ke-*k*:

$$\pi_k(\boldsymbol{x}) = P(Y=k|\boldsymbol{x}) = \frac{exp(g_k(\boldsymbol{x}))}{\sum_{k=1}^{q} exp(g_k(\boldsymbol{x}))} \tag{2.2}$$

di mana $\pi_k(\boldsymbol{x})$ menyatakan peluang kategori respon ke-*k* pada *p* peubah prediktor yang dinyatakan dalam vektor $\boldsymbol{x} = \left(1,\ x_1,\ x_2,\ x_3, \ldots,\ x_p\right)$ dan $g_1(\boldsymbol{x})$ = 0 karena kategori 1 digunakan sebagai pembanding, sehingga transformasi logit menjadi,

$$
\begin{aligned}
g_k(\boldsymbol{x}) &= ln\left(\frac{\pi_k(\boldsymbol{x})}{\pi_1(\boldsymbol{x})}\right)\\
&= ln\left(\frac{\frac{exp(g_k(x))}{\sum_{k=1}^{q} exp(g_k(x))}}{\frac{1}{\sum_{k=1}^{q} exp(g_k(x))}}\right)\\
&= ln\Big(exp\big(g_k(\boldsymbol{x})\big)\Big)\\
&= g_k(\boldsymbol{x})\\
&= \beta_{k0} + \beta_{k1}x_1 + \cdots + \beta_{kp}x_p \qquad (2.3)
\end{aligned}
$$

di mana:

$g_k(\boldsymbol{x})$: fungsi logit peubah respon untuk kategori ke-*k*

$\boldsymbol{x}$: vektor peubah prediktor

β_{kj} : koefisien model kategori ke-*k* peubah prediktor ke-*j*

x_j : nilai peubah prediktor ke-*j*; *j=1,2,...,p*

Pada regresi logistik dengan peubah respon tiga kategori (*k*=1,2,3) akan terbentuk dua fungsi logit. Karena *k*=1 sebagai kategori pembanding maka dua fungsi logit:

$$
\begin{aligned}
g_2(\boldsymbol{x}) &= ln\left(\frac{\pi_2(\boldsymbol{x})}{\pi_1(\boldsymbol{x})}\right)\\
&= ln\left(\frac{\frac{exp(g_2(x))}{\sum_{k=1}^{q} exp(g_k(x))}}{\frac{1}{\sum_{k=1}^{q} exp(g_k(x))}}\right)\\
&= ln\Big(exp\big(g_2(\boldsymbol{x})\big)\Big)\\
&= g_2(\boldsymbol{x})\\
&= \beta_{20} + \beta_{21}x_1 + \cdots + \beta_{2p}x_p \qquad (2.4)
\end{aligned}
$$

dan

$$
\begin{aligned}
g_3(\boldsymbol{x}) &= ln\left(\frac{\pi_3(\boldsymbol{x})}{\pi_1(\boldsymbol{x})}\right)\\
&= ln\left(\frac{\frac{exp(g_3(x))}{\sum_{k=1}^{q} exp(g_k(x))}}{\frac{1}{\sum_{k=1}^{q} exp(g_k(x))}}\right)\\
&= ln\Big(exp\big(g_3(\boldsymbol{x})\big)\Big)\\
&= g_3(\boldsymbol{x})\\
&= \beta_{30} + \beta_{31}x_1 + \cdots + \beta_{3p}x_p \qquad (2.5)
\end{aligned}
$$

kemudian membentuk model peluang bersyarat terhadap setiap kategori peubah respon:

$$\pi_1(\boldsymbol{x}) = P(Y=1|\boldsymbol{x}) = \frac{exp(g_1(\boldsymbol{x}))}{\sum_{k=1}^{3} exp(g_k(\boldsymbol{x}))} = \frac{1}{1+ exp(g_2(\boldsymbol{x}))+exp(g_3(\boldsymbol{x}))} \quad (2.6)$$

$$\pi_2(\boldsymbol{x}) = P(Y=2|\boldsymbol{x}) = \frac{exp(g_2(\boldsymbol{x}))}{\sum_{k=1}^{3} exp(g_k(\boldsymbol{x}))} = \frac{exp(g_2(\boldsymbol{x}))}{1+ exp(g_2(\boldsymbol{x}))+exp(g_3(\boldsymbol{x}))} \quad (2.7)$$

dan

$$\pi_3(\boldsymbol{x}) = P(Y=3|\boldsymbol{x}) = \frac{exp(g_3(\boldsymbol{x}))}{\sum_{k=1}^{3} exp(g_k(\boldsymbol{x}))} = \frac{exp(g_3(\boldsymbol{x}))}{1+ exp(g_2(\boldsymbol{x}))+exp(g_3(\boldsymbol{x}))} \quad (2.8)$$

akan didapatkan bentuk umum model peluang dengan 3 kategori respon:

$$\pi_k(\boldsymbol{x}) = P(Y=k|\boldsymbol{x}) = \frac{exp(g_k(\boldsymbol{x}))}{\sum_{k=1}^{3} exp(g_k(\boldsymbol{x}))} = \frac{exp(\beta_{k0}+\beta_{k1}x_1+\cdots+\beta_{kp}x_p)}{\sum_{k=1}^{3} exp(\beta_{k0}+\beta_{k1}x_1+\cdots+\beta_{kp}x_p)} \quad (2.9)$$

di mana $g_1(\boldsymbol{x})$ = 0 (Hosmer dan Lemeshow, 2000).

2.3. Struktur Data

Regresi logistik multinomial akan diterapkan pada data dengan struktur yang diperlihatkan pada Tabel 2.1 dan Tabel 2.2.

Tabel 2.1. Struktur Data untuk Peubah Prediktor

i	$X_{j,i}$				
	j				
	1	*2*	...	*p-1*	*p*
1	$x_{1,1}$	$x_{2,1}$	...	$x_{p-1,1}$	$x_{p,1}$
2	$x_{1,2}$	$x_{2,2}$	...	$x_{p-1,2}$	$x_{p,2}$
3	$x_{1,3}$	$x_{2,3}$	...	$x_{p-1,3}$	$x_{p,3}$
4	$x_{1,4}$	$x_{2,4}$	...	$x_{p-1,4}$	$x_{p,4}$
. . .	. . .	. . .	... . .	. . .	. . .
n	$x_{1,n}$	$x_{2,n}$	...	$x_{p-1,n}$	$x_{p,n}$

Tabel 2.2. Struktur Data untuk Peubah Respon

i	$Y_{k,i}$				
	k				
	1	*2*	…	*q-1*	*q*
1	$y_{1,1}$	$y_{2,1}$	⋯	$y_{q-1,1}$	$y_{q,1}$
2	$y_{1,2}$	$y_{2,2}$	⋯	$y_{q-1,2}$	$y_{q,2}$
3	$y_{1,3}$	$y_{2,3}$	⋯	$y_{q-1,3}$	$y_{q,3}$
4	$y_{1,4}$	$y_{2,4}$	⋯	$y_{q-1,4}$	$y_{q,4}$
. . .	. . .	. . .	⋯ . .	. . .	. . .
n	$y_{1,n}$	$y_{2,n}$	⋯	$y_{q-1,n}$	$y_{q,n}$

di mana:
$i = 1,2, \dots, n$; n : banyaknya pengamatan
$j = 1,2, \dots, p$; p : banyaknya peubah prediktor
$k = 1,2, \dots, q$; q : banyaknya kategori pada peubah respon

2.4. Pendugaan Parameter Regresi Logistik Multinomial

Pendugaan parameter β regresi logistik multinomial menggunakan metode kemungkinan maksimum. Menurut Hosmer dan Lemeshow (2000), untuk membuat fungsi *likelihood* model dengan kategori respon lebih dari dua (misal 3 kategori) terlebih dahulu membuat tiga peubah biner yang di beri kode 0 atau 1 untuk menandai anggota kelompok pada suatu pengamatan. Tiga peubah biner hanya untuk menjelaskan fungsi *likelihood* dan bukan digunakan untuk analisis regresi logistik multinomial yang sebenarnya.
Jika y =1 maka $y_{1i} = 1, y_{2i} = 0,$ dan $y_{3i} = 0.$
y =2 maka $y_{1i} = 0, y_{2i} = 1,$ dan $y_{3i} = 0.$
y =3 maka $y_{1i} = 0, y_{2i} = 0,$ dan $y_{3i} = 1.$
Fungsi *likelihood* untuk n contoh pengamatan saling bebas adalah:

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{n} [\pi_1(\boldsymbol{x})^{y_{1i}} \pi_2(\boldsymbol{x})^{y_{2i}} \pi_3(\boldsymbol{x})^{y_{3i}}]. \qquad (2.10)$$

di mana $\boldsymbol{\beta}$ = β_{k0}, β_{k1},…,β_{kp} dan y_{ki} = respon individu/pengamatan ke-*i* pada kategori ke-*k*.

Log *likelihood* persamaan 2.10 adalah:

$l(\boldsymbol{\beta}) = ln \prod_{i=1}^{n}[\pi_1(\boldsymbol{x})^{y_{1i}}\pi_2(\boldsymbol{x})^{y_{2i}}\pi_3(\boldsymbol{x})^{y_{3i}}]$

$= ln \prod_{i=1}^{n}[\left(\frac{1}{1+e^{g_2(x)}+e^{g_3(x)}}\right)^{y_{1i}} \left(\frac{e^{g_2(x)}}{1+e^{g_2(x)}+e^{g_3(x)}}\right)^{y_{2i}}$

$(\frac{e^{g_3(x)}}{1+e^{g_2(x)}+e^{g_3(x)}})^{y_{3i}}\ ]$

$= ln \prod_{i=1}^{n}[\left(e^{g_2(\boldsymbol{x})}\right)^{y_{2i}} \left(e^{g_3(\boldsymbol{x})}\right)^{y_{3i}} \left(\frac{1}{1+e^{g_2(x)}+e^{g_3(x)}}\right)^{y_{1i+}y_{2i+}y_{3i}}].$ (2.11)

dan diketahui bahwa $\sum_{i=1}^{n} y_{ki} = 1$, sehingga diperoleh:

$l(\boldsymbol{\beta}) =$

$$ln \prod_{i=1}^{n}[\left(e^{g_2(\boldsymbol{x})}\right)^{y_{2i}} \left(e^{g_3(\boldsymbol{x})}\right)^{y_{3i}} \left(\frac{1}{1+e^{g_2(\boldsymbol{x})}+e^{g_3(\boldsymbol{x})}}\right)^{y_{1i+}y_{2i+}y_{3i}}]$$

$= ln[\left(e^{g_2(\boldsymbol{x})}\right)^{\sum_{i=1}^{n} y_{2i}}\left(e^{g_3(\boldsymbol{x})}\right)^{\sum_{i=1}^{n} y_{3i}}\left(1+e^{g_2(\boldsymbol{x})}+e^{g_3(\boldsymbol{x})}\right)^{-\sum_{i=1}^{n}(y_{1i+}y_{2i+}y_{3i})}]$

$= ln\left(e^{g_2(\boldsymbol{x})}\right)^{\sum_{i=1}^{n} y_{2i}} + ln\left(e^{g_3(\boldsymbol{x})}\right)^{\sum_{i=1}^{n} y_{3i}} + ln\left(1+e^{g_2(\boldsymbol{x})}+e^{g_3(\boldsymbol{x})}\right)^{-\sum_{i=1}^{n}(y_{1i+}y_{2i+}y_{3i})}$

$= \sum_{i=1}^{n} y_{2i} ln(e^{g_2(\boldsymbol{x})}) + \sum_{i=1}^{n} y_{3i} ln(e^{g_3(\boldsymbol{x})}) - \sum_{i=1}^{n}(y_{1i+}y_{2i+}y_{3i}) ln(1 + e^{g_2(\boldsymbol{x})} + e^{g_3(\boldsymbol{x})})$

$= \sum_{i=1}^{n} y_{2i} ln(e^{g_2(\boldsymbol{x})}) + \sum_{i=1}^{n} y_{3i} ln(e^{g_3(\boldsymbol{x})}) - ln(1 + e^{g_2(\boldsymbol{x})} + e^{g_3(\boldsymbol{x})})).$

$= \sum_{i=1}^{n}[[y_{2i}\, g_2(\boldsymbol{x}) + y_{3i}(\boldsymbol{x_i}) g_3(\boldsymbol{x})] - ln(1 + e^{g_2(\boldsymbol{x})} + e^{g_3(\boldsymbol{x})})].$ (2.12)

Pandang fungsi skor *likelihood* ($\boldsymbol{u}$) sebagai turunan pertama $l(\boldsymbol{\beta})$ terhadap β_{kj}. Untuk menyederhanakan notasi, π_{ki}=$\pi_k(\boldsymbol{x}_i)$.

$$\frac{\partial l(\boldsymbol{\beta})}{\partial \beta_{kj}} = 0$$

$$\frac{\partial l(\boldsymbol{\beta})}{\partial \beta_{kj}} = \sum_{i=1}^{n}\Big(y_{2i}\left(x_{0i}+x_{1i}+x_{2i}+\cdots+x_{pi}\right) + y_{3i}\left(x_{0i}+x_{1i}+x_{2i}+\cdots+x_{pi}\right)\Big)$$

$$-ln\left(1+e^{g_2(\boldsymbol{x})}+e^{g_3(\boldsymbol{x})}\right)$$

$$= \sum_{i=1}^{n} x_{ji}y_{ki} - \frac{1}{(1+e^{g_2(\boldsymbol{x})}+e^{g_3(\boldsymbol{x})}))}$$

$$\left(1+e^{g_2(\boldsymbol{x})}+e^{g_3(\boldsymbol{x})})\right)^{(\beta_{kj})}$$

$= \sum_{i=1}^{n} x_{ji}y_{ki} - \frac{1}{(1+e^{g_2(x)}+e^{g_3(x)})}$

$$\left[e^{g_2(\boldsymbol{x})(x_{0i}+x_{1i}+x_{2i}+\cdots+x_{pi})} + e^{g_3(\boldsymbol{x})(x_{0i}+x_{1i}+x_{2i}+\cdots+x_{pi})}\right]$$
$$= \sum_{i=1}^{n} x_{ji}y_{ki} - \sum_{i=1}^{n} x_{ji}\left[\frac{(e^{g_2(\boldsymbol{x})} + e^{g_3(\boldsymbol{x})})}{1 + e^{g_2(\boldsymbol{x})} + e^{g_3(\boldsymbol{x})})}\right] = 0$$

akan diperoleh persamaan skor *likelihood* ($\boldsymbol{u}$):

$$\sum_{i=1}^{n} x_{ji}y_{ki} - \sum_{i=1}^{n} x_{ji}\pi_{ki} = 0$$
$$\sum_{i=1}^{n} x_{ji}(y_{ki} - \pi_{ki}) = 0$$
$$\boldsymbol{u} = \frac{\partial l(\boldsymbol{\beta})}{\partial \beta_{kj}} = \sum_{i=1}^{n} x_{ji}(y_{ki} - \pi_{ki}) \qquad (2.13)$$

di mana:

x_{ji} : pengamatan ke-*i*, peubah prediktor ke-*j*.

y_{ki} : pengamatan ke-*i*, peubah respon kategori ke-*k*.

π_{ki} : peluang kategori respon ke-*k* pada pengamatan ke-*i*.

Sebagaimana diketahui bahwa proses pendugaan parameter dengan metode MLE didasarkan pada turunan parsial fungsi *likelihood* terhadap parameter. Akan tetapi, turunan parsial pertama fungsi log *likelihood* terhadap parameter merupakan fungsi non-linier. Oleh karena itu, diperlukan metode iteratif *Newton-Raphson* untuk mendapatkan solusi persamaan nonlinier.

Iterasi *Newton-Raphson* adalah:

$$\widehat{\boldsymbol{\beta}}^{(t+1)} = \boldsymbol{\beta}^{(t)} - [\boldsymbol{H}^{(t)}]^{-1}\boldsymbol{g}^{(t)} \qquad (2.14)$$

di mana t adalah proses iterasi (1,2,3,...) dan nilai $\boldsymbol{H}^{(t)}$ dan $\boldsymbol{g}^{(t)}$ berupa matriks:

$$\boldsymbol{H}^{(t)} = \frac{\partial^2 l(\boldsymbol{\beta})}{\partial \beta_{kj}\partial \beta_{kj}} = \boldsymbol{X'V^{(t)}X} \qquad (2.15)$$
$$\boldsymbol{g}^{(t)} = \frac{\partial l(\boldsymbol{\beta})}{\partial \beta_{kj}} = \sum_{i=1}^{n} x_{ji}(y_{ki} - \pi_{ki}) = \boldsymbol{X'}(\boldsymbol{Y} - \boldsymbol{\pi}^{(t)}) \qquad (2.16)$$

$\boldsymbol{H}^{(t)}$ merupakan matriks turunan parsial ke-2 fungsi log *likelihood* terhadap parameter sedangkan $\boldsymbol{g}^{(t)}$ adalah turunan parsial pertama fungsi log *likelihood* terhadap parameter. $[\boldsymbol{H}^{(t)}]^{-1}$ merupakan matriks peragam (Agresti, 2002). Prosedur iterasi *Newton-Raphson* adalah:

1. Menentukan nilai awal penduga parameter yaitu $\boldsymbol{\beta}^{(t)}$. Nilai $\boldsymbol{\beta}^{(t)}$ berdekatan dengan selisih antara nilai penduga dan penduga yang

diharapkan, hal ini disebabkan agar langkah iterasi tidak terlalu lama, untuk mempermudah perhitungan digunakan 0 sebagai nilai penduga awal iterasi.

2. Membentuk data matriks

$$\boldsymbol{X}_{n\times(p+1)} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & & x_{2p} \\ & \vdots & & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \text{ dan } \boldsymbol{Y}_{n\times 1} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

3. Menentukan model regresi logistik

$$g_k(\boldsymbol{x}) = \beta_{k0} + \beta_{k1}x_1 + \cdots + \beta_{kp}x_p$$

4. Menentukan matriks $\boldsymbol{V^{(t)}}$

$$\boldsymbol{V^{(t)}}_{n\times n} = \begin{bmatrix} \pi_1{}^{(t)}(1-\pi_1{}^{(t)}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \pi_n{}^{(t)}(1-\pi_n{}^{(t)}) \end{bmatrix}$$

5. Mensubstitusi hasil persamaan (2.14) dengan

$$\boldsymbol{H}^{(t)}{}_{(p+1)\times(p+1)} = \frac{\partial^2 l(\boldsymbol{\beta})}{\partial\beta_{kj}\partial\beta_{kj}} = \boldsymbol{X'V^{(t)}X}$$

$$= \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & x_{31} & & x_{n1} \\ & \vdots & & \ddots & \vdots \\ x_{1p} & x_{2p} & x_{3p} & \cdots & x_{np} \end{bmatrix}' \begin{bmatrix} \pi_1{}^{(t)}(1-\pi_1{}^{(t)}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \pi_n{}^{(t)}(1-\pi_n{}^{(t)}) \end{bmatrix}$$

$$\begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & & x_{2p} \\ & \vdots & & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

dan

$$\boldsymbol{g}^{(t)} = \boldsymbol{X}'(\boldsymbol{Y} - \boldsymbol{\pi^{(t)}})$$

$$\boldsymbol{\pi^{(t)}}_{n\times 1} = \begin{bmatrix} \pi_1{}^{(t)} \\ \pi_2{}^{(t)} \\ \vdots \\ \pi_n{}^{(t)} \end{bmatrix}$$

6. Diperoleh $\widehat{\boldsymbol{\beta}}^{(t+1)}$

7. Menghitung selisih $(\widehat{\boldsymbol{\beta}}^{(t+1)} - \boldsymbol{\beta}^{(t)})$, jika selisih < 1×10^{-6} maka $\widehat{\boldsymbol{\beta}}^{(t+1)}$ merupakan hasil pendugaan dan proses iterasi diulang hingga diperoleh nilai $\boldsymbol{\beta}$ yang konvergen. Menurut Agresti (2002)

batas konvergen adalah $\left|\widehat{\boldsymbol{\beta}}^{(t+1)} - \boldsymbol{\beta}^{(t)}\right| < \varepsilon$, di mana $\varepsilon = 1 \times 10^{-6}$.

2.5. Pemisahan dalam Metode *Maximum Likelihood Estimation*

2.5.1. Pemisahan Sempurna dan Pemisahan Kurang Sempurna

Parameter tak hingga didapat dari persamaan *likelihood* yang tidak memiliki solusi hingga, dengan kata lain tidak ada penduga *maximum likelihood*. Hal ini diakibatkan oleh konfigurasi titik contoh hasil pengamatan yang disebut pemisahan sempurna dan pemisahan kurang sempurna (Albert dan Anderson, 1984).

Pandang y_i sebagai respon multinomial pengamatan ke-*i* dan $\boldsymbol{x_i}' = \left(1,\ x_{i1},\ x_{i2},\ x_{i3}, \ldots,\ x_{ip}\right)$ berukuran 1×(p+1) sebagai vektor dari nilai peubah prediktor dan konstanta pengamatan ke-*i* (*i*= *1, 2, ..., n*), sehingga terdapat penambahan unsur 1 pada $\boldsymbol{x_i}'$.

Pemisahan sempurna terjadi apabila terdapat vektor $\boldsymbol{b}_{(p+1)x1}$ yang secara tepat meramalkan semua respon y_i, sehingga:

$$\boldsymbol{b}'\boldsymbol{x_i} = 0 \qquad \text{,untuk } y_i = 1, 2 \text{ dan } 3 \tag{2.17}$$

di mana $\boldsymbol{b}' = \left(b_0,\ b_1,\ b_2,\ \ldots, b_p\right)$ adalah vektor koefisien dari fungsi linier peubah prediktor ($f(x_0,\ x_1, \ldots, x_p)$) sama dengan nol yang mengakibatkan hasil pengamatan terpisah sempurna.

$f(x_1, x_2, \ldots, x_p) = 0$

$b_0 + b_1x_1 + b_2x_2 + \ldots + b_px_p = 0$

$$\left(b_0,\quad b_1,\quad b_2,\quad \ldots,\quad b_p\right)\begin{pmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} = 0 \tag{2.18}$$

Fungsi $f(x_1,\ x_2, \ldots, x_p)$ diperoleh melalui diagram pencar di mana peubah prediktor merupakan sumbu koordinat. Pilih sembarang titik p yang memisahkan respon pengamatan secara sempurna.

Berikut adalah data hipotetik berukuran 10 yang mengilustrasikan keadaan pemisahan sempurna di mana Y adalah peubah respon dan peubah prediktor X_1, X_2 bersifat kontinu yang menghasilkan diagram pencar seperti tampak dalam Gambar 2.1.

Tabel 2.3. Data Hipotetik 1 Pemisahan Sempurna

i	y_i	x_{i1}	x_{i2}
1	1	56	100
2	1	43	110
3	1	52	110
4	1	60	110
5	1	46	100
6	2	48	160
7	2	41	140
8	2	50	150
9	2	55	150
10	2	49	140

Misal dipilih dua pasang sembarang titik $(x_{i1},\ x_{i2})$ yang tidak disajikan di Tabel 2.3 yang memisahkan hasil pengamatan secara sempurna yakni (42,120) dan (51,130) di mana $0 \leq x_{i1} \leq \infty$ dan $110 \leq x_{i2} \leq 140$. Fungsi linier yang melewati kedua titik adalah:

$$\frac{x_1 - 42}{51 - 42} = \frac{x_2 - 120}{130 - 120}$$

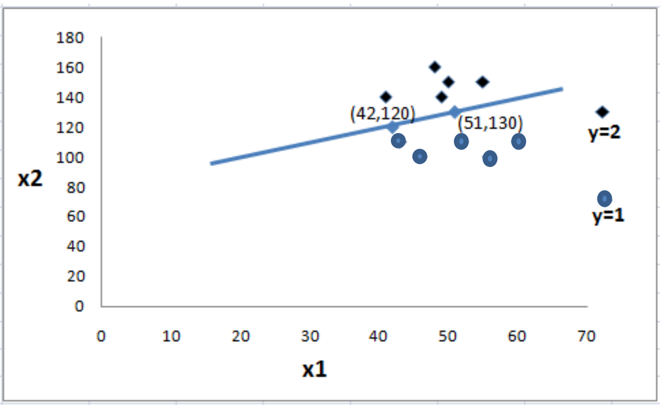
$$\frac{x_1 - 42}{9} = \frac{x_2 - 120}{10}$$

$$10(x_1 - 42) = 9(x_2 - 120)$$

$$10x_1 - 420 = 9x_2 - 1080$$

$$660 + 10x_1 - 9x_2 = 0$$

$$10x_1 - 9x_2 = -660 \qquad (2.19)$$



Gambar 2.1. Diagram Pencar Data Hipotetik 1

Grafik memperlihatkan bahwa fungsi $f(x_1,x_2)$ = 660+10$x_1 - 9x_2$ = 0 secara sempurna memisahkan semua pengamatan ke dalam kategori respon masing-masing, sehingga pengamatan dengan respon yang sama berada pada sisi yang sama. Vektor $\boldsymbol{b}$ yang memenuhi persamaan 2.17 adalah $\boldsymbol{b}'$ = (660,10,-9).

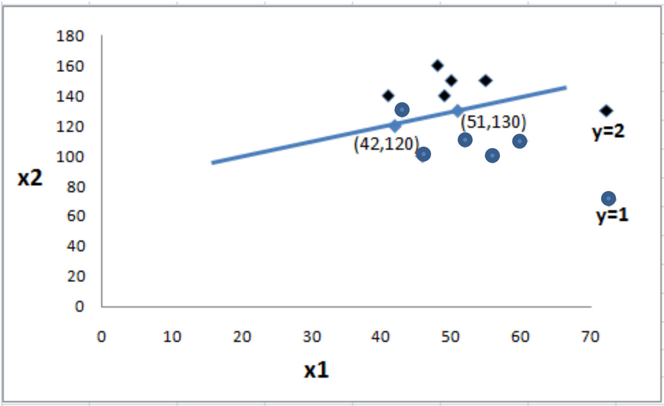
Data ini dimodelkan dengan model regresi logistik sehingga pendugaan parameter menggunakan metode MLE dan metode iteratif *Newton-Raphson*.

Pemisahan kurang sempurna terjadi ketika terdapat vektor $\boldsymbol{b}$ yang mengakibatkan:

$$\begin{cases} \boldsymbol{b}'\boldsymbol{x}_i \geq 0 & , y_i = 1 \\ \boldsymbol{b}'\boldsymbol{x}_i = 0 & , y_i = 2 \\ \boldsymbol{b}'\boldsymbol{x}_i \leq 0 & , y_i = 3 \end{cases} \tag{2.20}$$

ketika paling tidak terdapat satu i yang memenuhi (2.20)

Vektor $\boldsymbol{b}$ pada pemisahan kurang sempurna diperoleh dari fungsi linier peubah prediktor yang melewati paling sedikit satu pengamatan pada setiap kategori respon. Modifikasi dilakukan pada data hipotetik 1 untuk menghasilkan kondisi pemisahan kurang sempurna di mana x_{22} = 110 diubah menjadi x_{22} = 130 (Lusiana, 2012).



Gambar 2.2. Diagram Pencar Modifikasi Data Hipotetik 1

Tampak pada Gambar 2.2 bahwa titik modifikasi (43,130) pada pengamatan ke-2 berada di daerah y=2. Pada awalnya titik awal pengamatan kedua (43,110) berada di kawasan y=1 namun ketika titik tersebut dimodifikasi maka mengakibatkan pemisahan kurang sempurna.

2.5.2. Pemeriksaan Keberadaan Pemisahan

Penentuan vektor ***b*** yang mengindikasi keberadaan pemisahan sempurna dan kurang sempurna akan sulit dilakukan jika model mengandung banyak peubah prediktor (p>2). Oleh karena itu dibutuhkan metode *Maximum Likelihood Estimation* dan metode iteratif *Newton-Raphson* untuk memeriksa keberadaan pemisahan. Berikut ini adalah prosedur pemeriksaan pemisahan yang diusulkan oleh Albert dan Anderson (1984) dalam Allison (2008):

a. Apabila dalam iterasi ke delapan ($t \leq 8$) kriteria kekonvergenan terpenuhi, maka dapat dikatakan bahwa tidak terjadi pemisahan.
b. Untuk iterasi lebih dari delapan ($t > 8$), hitung peluang ketepatan alokasi setiap pengamatan ke-*i*:

$$P(\hat{Y}_i = y_i) = \frac{1}{1 + exp[(2y_i - 1)\hat{\boldsymbol{\beta}}\boldsymbol{x}_i]} \quad , y_i = 1, 2, 3 \qquad (2.21)$$

Apabila peluang ketepatan alokasi sama dengan 1 untuk semua pengamatan maka terjadi pemisahan sempurna dan proses iterasi dihentikan.
c. Jika peluang ketepatan alokasi terbesar lebih dari 0.95 untuk beberapa pengamatan, maka periksa ragam penduga parameter yang dibakukan bagi iterasi tersebut. Apabila ragam penduga lebih besar dari 5, dikatakan terjadi pemisahan kurang sempurna dan proses iterasi dihentikan.

2.6. Metode *Penalized Maximum Likelihood Estimation*

Prinsip dasar metode PMLE adalah memodifikasi fungsi skor *likelihood* menjadi fungsi skor *penalized likelihood*, $\boldsymbol{u}^*$ (Heinze dan Schemper, 2002):

$$\frac{\partial(l*(\boldsymbol{\beta}))}{\partial\beta_{kj}} = \boldsymbol{u}^* = \sum_i\{(y_{ki} + \frac{t_i}{2}) - \pi_k(\boldsymbol{x}_i)(1+t_i)\}\, x_{ji} \qquad (2.22)$$

di mana $l * (\boldsymbol{\beta}) = l(\boldsymbol{\beta}) + ½ \ln |I(\boldsymbol{\beta})|$ adalah log-likelihood modifikasi dan $|I(\boldsymbol{\beta})|^{1/2}$ disebut *Jeffrey's invariant prior.*

McCullagh dan Nelder (1989) menyatakan bentuk umum vektor bias $O(N^{-1})$, yang diusulkan oleh Firth pada tahun 1993 yaitu suatu metode untuk menghilangkan bias orde pertama, $O(N^{-1})$:

$$b(\boldsymbol{\beta}) = (\boldsymbol{X'WX})^{-1}\, \boldsymbol{X'W\xi} = (I(\boldsymbol{\beta}))^{-1}\, \boldsymbol{X'W\xi} \qquad (2.23)$$

di mana:

$b(\boldsymbol{\beta})$: bias orde pertama $O(N^{-1})$
$I(\boldsymbol{\beta})$: matriks informasi Fisher terevaluasi pada $\boldsymbol{\beta}$
$\boldsymbol{X}$: matriks rancangan prediktor berukuran n x (p+1)

$\boldsymbol{W}$: matriks diag $[\pi_k(\boldsymbol{x}_i)(1- \pi_k(\boldsymbol{x}_i))])$.

$\boldsymbol{W\xi}$: matriks diagonal dengan unsur $t_i(\pi_k(\boldsymbol{x}_i) - \frac{1}{2})$

t_i : unsur diagonal ke-*i* matriks topi $\boldsymbol{T}$

$\boldsymbol{T}$ $= \boldsymbol{W^{1/2}X(X'WX)^{-1}X'\,W^{1/2}}$

i $= 1, 2, ..., n$

Pandang $l(\boldsymbol{\beta})$ pada persamaan (2.12) dan skor *likelihood* ($\boldsymbol{u}$) pada persamaan (2.13). Firth (1993) mengusulkan bahwa bias orde pertama $O(N^{-1})$ dapat dihilangkan dengan memodifikasi $\boldsymbol{u}$:

$$\boldsymbol{u}^* = \boldsymbol{u} - I(\boldsymbol{\beta})b(\boldsymbol{\beta})$$
$$= \boldsymbol{u} - I(\boldsymbol{\beta})(I(\boldsymbol{\beta}))^{-1} \boldsymbol{X'W\xi}$$
$$= \boldsymbol{u} - \boldsymbol{X'W\xi} \tag{2.24}$$

dengan demikian unsur-unsur $\boldsymbol{u}^*$,

$$u_{kj}^* = \sum_i(y_{ki} - \pi_k(\boldsymbol{x}_i))x_{ji} - \sum_i(t_i(\pi_k(\boldsymbol{x}_i) - \tfrac{1}{2})\, x_{ji}$$
$$= \sum_i(y_{ki} - \pi_k(\boldsymbol{x}_i) - t_i(\pi_k(\boldsymbol{x}_i) - \tfrac{1}{2})\, x_{ji}$$
$$= \sum_i(y_{ki} - \pi_k(\boldsymbol{x}_i) - t_i\pi_k(\boldsymbol{x}_i) + \tfrac{t_i}{2})\, x_{ji}$$
$$= \sum_i\{(y_{ki} + \tfrac{t_i}{2}) - \pi_k(\boldsymbol{x}_i)(1 + t_i)\}x_{ji} \tag{2.25}$$

Fungsi skor *likelihood* modifikasi ini bersesuaian dengan fungsi *penalized likelihood*, $L^*(\boldsymbol{\beta})$:

$$L^*(\boldsymbol{\beta}) = L(\boldsymbol{\beta})\, |I(\boldsymbol{\beta})|^{1/2} \tag{2.26}$$

dan *penalized log-likelihood*, $l^*(\boldsymbol{\beta})$:

$$l^*(\boldsymbol{\beta}) = l(\boldsymbol{\beta}) + \tfrac{1}{2}\ln|\, I(\boldsymbol{\beta})| \tag{2.27}$$

di mana:

$L(\boldsymbol{\beta})$: *Likelihood* tanpa modifikasi

$L^*(\boldsymbol{\beta})$: *Likelihood* modifikasi

$l(\boldsymbol{\beta})$: *log-likelihood* tanpa modifikasi

$l^*(\boldsymbol{\beta})$: *log-likelihood* modifikasi

$|I(\boldsymbol{\beta})|^{1/2}$: *Jeffrey's invariant prior*

Penduga parameter bagi model regresi logistik multinomial diperoleh melalui cara yang sama seperti metode MLE yakni dengan memaksimumkan $l^*(\boldsymbol{\beta})$. *Jeffrey's invariant prior* adalah prior noninformatif yang menjamin bahwa suatu parameter akan selalu memiliki prior yang sama meskipun telah mengalami transformasi (Hewson, 2009).

Firth (1993) berpendapat bahwa dengan jaminan $\boldsymbol{X}$ berpangkat penuh, maka $\ln|I(\boldsymbol{\beta})|$ akan bersifat konkaf dan *unbounded* *below* ketika $\beta \to \infty$. Fungsi *log-likelihood* $l(\boldsymbol{\beta})$ bersifat konkaf dan *bounded above*. Hal ini mengakibatkan metode PMLE akan selalu menghasilkan penduga unik dan berhingga.

2.7. Pengujian Parameter

Pengujian terhadap *p* parameter dilakukan secara parsial dan simultan untuk mengetahui apakah peubah prediktor yang terdapat dalam model berpengaruh terhadap peubah respon.

1. Uji Simultan

Digunakan statistik uji *G* untuk menguji koefisien regresi secara serentak. Pengujian ini merupakan rasio kemungkinan maksimum, berdasarkan hipotesis (Hosmer & Lemeshow, 2000):

$\mathrm{H}_0 : \beta_{kj} = 0$; peubah prediktor tidak berpengaruh terhadap peubah respon.

$\mathrm{H}_1 : \beta_{kj} \neq 0$; paling tidak terdapat satu peubah prediktor yang berpengaruh terhadap peubah respon.

Jika H_0 benar, statistik uji *G*:

$$-2\ln\left[\frac{L_0}{L_1}\right] \sim \chi^2_{(p)} \tag{2.28}$$

di mana:

L_0 : log *likelihood* tanpa peubah prediktor.

$L_0 = \left(\frac{n_0}{n}\right)^{n_o}\left(\frac{n_1}{n}\right)^{n_1}\ldots\left(\frac{n_k}{n}\right)^{n_k}$; n_k = banyaknya individu pada kategori ke-*k*. (2.29)

L_1 : log *likelihood* model intersep dan semua peubah prediktor.

$$L_1 = \prod_{i=1}^{n}\left[\pi_0(\boldsymbol{x_i})^{y_{0i}}\pi_1(\boldsymbol{x_i})^{y_{1i}}\ldots\pi_{k-1}(\boldsymbol{x_i})^{y_{(k-1)i}}\right] \tag{2.30}$$

di mana:

$\pi_k(\boldsymbol{x_i})$: peluang kategori respon ke-*k* pada pengamatan ke-*i*.

y_{ki} : anggota peubah respon pada kategori ke-*k* pengamatan ke-*i*.

Tolak H_0 jika $G > \chi^2_{(\alpha,p)}$.

2. Uji Parsial

Pengujian parameter secara parsial dilakukan untuk menguji keberartian parameter secara individual. Hasil pengujian secara parsial akan menunjukkan apakah suatu peubah prediktor layak masuk dalam model. Pengujian ini menggunakan statistik uji *Wald*, berdasarkan hipotesis (Hosmer & Lemeshow, 2000):

$\mathrm{H}_0 : \beta_{kj} = 0$; peubah prediktor ke-*j* pada kategori ke-*k* tidak berpengaruh terhadap peubah respon.

H_1 : $\beta_{kj} \neq 0$; peubah prediktor ke-*j* pada kategori ke-*k* berpengaruh terhadap peubah respon.

Sebaran penarikan contoh bagi $\hat{\beta}_{kj}$ adalah

$\hat{\beta}_{kj} \sim N(\beta_{kj}, Var(\hat{\beta}_{kj}))$

Jika H_0 benar, statistik uji *Wald* (W_{kj}):

$$\frac{\hat{\beta}_{kj}}{SE(\hat{\beta}_{kj})} \sim \mathrm{N}(0,1) \qquad (2.31)$$

di mana:

$SE(\hat{\beta}_{kj})$ = salah baku penduga β_{kj} dan merupakan akar diagonal utama matriks peragam.

$$SE(\hat{\beta}_{kj}) = \sqrt{var(\hat{\beta}_{kj})}$$

$$var(\hat{\beta}_{kj}) = diag[\boldsymbol{X'VX}]^{-1}$$

di mana $\boldsymbol{X}$ adalah matriks berukuran $n \times (p+1)$ dan $\boldsymbol{V}$ merupakan matriks diagonal $\hat{\pi}_i(1-\hat{\pi}_i)$ yang berukuran $n \times n$.

$$\boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{p1} \\ 1 & x_{12} & \cdots & x_{p2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \cdots & x_{pn} \end{bmatrix}$$

$$\boldsymbol{V} = \begin{bmatrix} \hat{\pi}_1(1-\hat{\pi}_1) & 0 & \cdots & 0 \\ 0 & \hat{\pi}_2(1-\hat{\pi}_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{\pi}_p(1-\hat{\pi}_p) \end{bmatrix}$$

Tolak H_0 jika $|W| > Z_{\alpha/2}$.

2.8. Pemilihan Model Terbaik

Salah satu metode untuk memilih model regresi logistik terbaik apabila peubah prediktor tidak nyata menurut uji parsial adalah metode *backward* yang diawali dengan mengikutsertakan semua peubah prediktor, kemudian secara bertahap mereduksi peubah yang memiliki statistik uji *Wald* (W_{kj}) mutlak terkecil (Agresti, 2002). Proses dihentikan ketika pereduksian peubah prediktor mengakibatkan model tidak sesuai. Misal dibutuhkan *m* reduksi berlandaskan hipotesis:

H_0 : model sesuai

H_1 : model tidak sesuai

Jika H_0 benar, statistik uji G_r^2:

$$G_{p-r+1} - G_{p-r} \sim X^2_{(1)} \tag{2.32}$$

di mana:

r = 1, 2,..., m ; m : banyaknya reduksi

p : banyaknya peubah prediktor

G_r^2 : statistik uji G^2 pada pereduksian ke-r

G_{p-r+1} : statistik uji G sebelum pereduksian ke-r

$$= -2\ln[\frac{L_0(\boldsymbol{\beta})}{L_{p-r+1}(\boldsymbol{\beta})}] \tag{2.33}$$

G_{p-r} : statistik uji G setelah pereduksian ke-r

$$= -2\ln[\frac{L_0(\boldsymbol{\beta})}{L_{p-r}(\boldsymbol{\beta})}] \tag{2.34}$$

Tolak H_0 jika $G_r^2 > \chi^2_{(\alpha,1)}$.

2.9. Pengujian Kelayakan Model

Pengujian kelayakan model dilakukan untuk mengetahui apakah model yang diperoleh layak, berdasarkan hipotesis :

H_0 : model layak

H_1 : model tidak layak

Jika H_0 benar, statistik uji $\chi^2_{Pearson}$:

$$\sum_{k=1}^{q=3} \frac{(y_{ki} - n_k\pi_k(\boldsymbol{x_i}))^2}{n_k\pi_k(\boldsymbol{x_i})(1-\pi_k(\boldsymbol{x_i}))} \sim \chi^2_{(n-p-1)} \tag{2.35}$$

di mana:

y_{ki} : nilai peubah respon pada kategori ke-k.

n_k : banyaknya pengamatan pada kategori ke-k.

$\pi_k(\boldsymbol{x_i})$: peluang kategori respon ke-k pada pengamatanke-i.

Tolak H_0 jika $\chi^2_{pearson} > \chi^2_{(\alpha,n-p-1)}$.

2.10. Interpretasi Model

Menurut Hosmer dan Lemeshow (2000), *odds ratio* (OR) merupakan suatu ukuran untuk mengetahui tingkat risiko hubungan suatu peubah prediktor terhadap peubah respon. *Odds ratio* digunakan untuk melakukan interpretasi terhadap model yang sebelumnya dilakukan pengujian signifikansi parameter model. *Odds Ratio* yang terbentuk adalah sebanyak fungsi logit yang didapatkan. Untuk nilai $x = 1$ dan $x = 0$, maka *odds ratio*:

$$\widehat{OR} = \frac{exp\left(\beta_{k0} + \beta_{kj}(1)\right)}{exp\left(\beta_{k0} + \beta_{kj}(0)\right)}$$

$$= exp(\beta_{kj}) \tag{2.36}$$

di mana β_{kj} adalah koefisien model regresi kategori ke-k peubah prediktor ke-j.

Odds ratio (OR) pada kategori $y = k$ merupakan perbandingan antara $x = a$ dan $x = b$, di mana *P(y = 2* dan *3*) adalah peluang kejadian sukses dan *P(y = 1)* adalah peluang kejadian gagal. Regresi logistik multinomial pada respon dengan tiga kategori diberi kode 1, 2 dan 3 akan menghasilkan dua *odds ratio* yaitu;

1. Perbandingan kategori 2 dan 1.

$$\widehat{OR}_1 = \frac{(y=2|x=a)/(y=1|x=a)}{(y=2|x=b)/(y=1|x=b)} \tag{2.37}$$

2. Perbandingan kategori 3 dan 1.

$$\widehat{OR}_2 = \frac{(y=3|x=a)/(y=1|x=a)}{(y=3|x=b)/(y=1|x=b)} \tag{2.38}$$

Odds Ratio selalu bernilai positif. Misal nilai $\widehat{OR}$ = 1 berarti bahwa $x = a$ mempunyai risiko yang sama dengan $x = b$ untuk menghasilkan $y = k$. Apabila 1 < $\widehat{OR}$, dapat dikatakan bahwa $x = a$ memiliki risiko lebih tinggi $\widehat{OR}$ kali dibandingkan $x = b$ untuk menghasilkan $y = k$. Sebaliknya, jika 0 < $\widehat{OR}$< 1 dapat dikatakan bahwa $x = a$ memiliki risiko lebih rendah $\widehat{OR}$ kali dibandingkan $x = b$ untuk menghasilkan $y = k$.

2.11. Tinjauan Non Statistika

Penyakit jantung koroner (PJK) adalah kelainan pada satu atau lebih pembuluh arteri koroner di mana terdapat penebalan dinding dalam pembuluh darah disertai aterosklerosis yang akan mempersempit limen arteri koroner dan akan mengganggu aliran darah ke otot jantung sehingga terjadi kerusakan dan gangguan pada otot jantung (Hariadi dan Arsad, 2005).

Banyak faktor yang mempengaruhi terjadi PJK, sehingga usaha pencegahan sangat diperlukan. Usaha pencegahan ini harus dilakukan sedapat mungkin dengan cara pengendalian faktor-faktor risiko PJK (Djohan, 2004).

Ada faktor risiko PJK yang tidak dapat dikendalikan, ada pula yang dapat dikendalikan. Faktor risiko yang tidak dapat dikendalikan adalah usia sedangkan yang dapat dikendalikan adalah tekanan darah dan profil lemak darah meliputi kadar kolesterol, HDL, LDL dan trigliserid (Wibowo, 2001).

Penjelasan profil lemak darah adalah:

a. Kadar Kolesterol
   Seseorang dengan kadar kolesterol tinggi (250-275 mg/dl) memiliki risiko menderita penyakit jantung lebih dari dua kali lipat daripada seseorang dengan kadar kolesterol kurang dari 175 mg/dl. Kenaikan kadar kolesterol berkaitan dengan peningkatan kadar LDL dalam darah.
b. Kadar LDL (Low Density Lipoprotein)
   Kadar LDL dalam darah yang tinggi berperan dalam timbulnya aterosklerosis, yaitu terhambatnya aliran darah koroner karena menumpuknya partikel kolesterol LDL di sepanjang dinding pembuluh darah yang membuat diameter pembuluh darah menyempit dan aliran darah terganggu sehingga akan menurunkan suplai oksigen. Kadar LDL yang dianggap tinggi adalah lebih dari 160 mg/dl.
c. Kadar HDL (High Density Lipoprotein)
   HDL dikenal sebagai kolesterol baik. HDL memiliki kemampuan untuk membersihkan tumpukan lemak yang menempel pada dinding pembuluh darah. Sehingga kadar HDL yang tinggi memiliki efek proteksi bagi jantung. Sebaliknya penurunan kadar HDL akan menyebabkan penumpukkan kolesterol di jaringan. Kadar HDL yang dianggap rendah adalah kurang dari 35 mg/dl.
d. Kadar Trigliserid
   Kadar trigliserid yang tinggi akan meningkatkan kadar LDL dan menurunkan kadar HDL. Patokan ideal kadar trigliserid adalah kurang dari 150 mg/dl.
   (Majid, 2010).

Penyakit jantung koroner sudah umum disertai hipertensi atau disertai hipertensi dan diabetes mellitus. Hipertensi merupakan penyakit multifaktorial yang muncul karena interaksi berbagai faktor. Dengan bertambah umur, maka tekanan darah akan meningkat karena pada usia yang lanjut dinding arteri akan mengalami penebalan yang disebabkan ada penumpukkan zat kolagen pada lapisan otot sehingga pembuluh darah akan berangsur-angsur menyempit dan menjadi kaku (Adhita , 2010).

Diabetes mellitus adalah penyakit yang ditandai peningkatan kadar gula darah yang diakibatkan oleh kerja hormon insulin. Apabila organ tubuh sudah tidak berfungsi dengan baik seperti tidak terkendalinya tekanan darah atau kadar profil lemak darah maka akan mengganggu kerja hormon. Apabila hormon insulin terganggu, akan terjadi kenaikan kadar gula dalam darah dan bila hal ini bertahan dalam waktu yang lama maka akan memicu terjadinya kerusakan serius pada beberapa sistem organ yakni syaraf dan pembuluh darah (Adhita , 2010).

Penyakit kardiovaskuler seperti penyakit jantung koroner merupakan penyakit kematian yang utama pada seseorang dengan diabetes mellitus dan disertai penyakit lain seperti hipertensi (Adhita , 2010).

BAB III
METODE PENELITIAN

3.1. Sumber Data

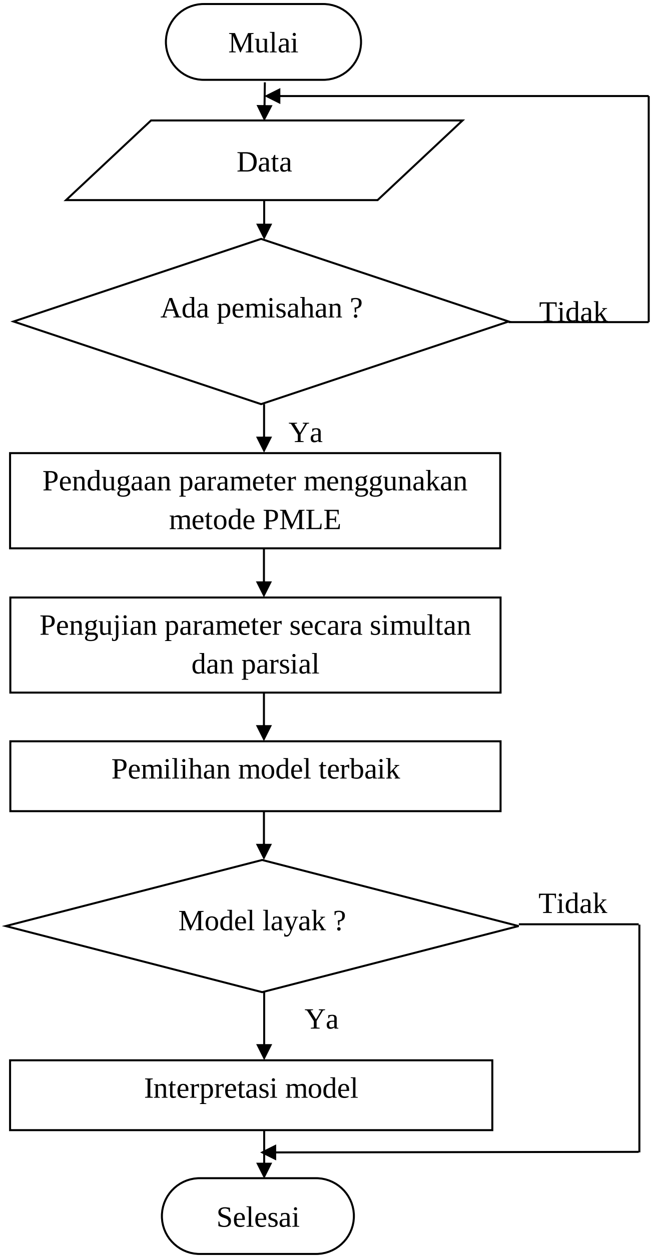
Data sekunder bersumber pada hasil penelitian yang dilakukan oleh Wibowo (2001) tentang profil pasien poli kardiologi RSUD Saiful Anwar Kota Malang. Peubah respon adalah tingkatan Penyakit Jantung Koroner (PJK) terdiri dari tiga kategori yaitu penyakit jantung koroner (1), PJK disertai hipertensi (2) dan PJK disertai hipertensi dan diabetes mellitus (3) berdasarkan 7 peubah prediktor:

X_1 : Usia (tahun)
X_2 : Tekanan darah *sistole* (mmHg)
X_3 : Tekanan darah *diastole* (mmHg)
X_4 : Kolesterol (mg/dl)
X_5 : HDL (mg/dl)
X_6 : LDL (mg/dl)
X_7 : Trigliserid (mg/dl)

3.2. Prosedur Analisis Data

1. Mendeskripsikan profil pasien poli kardiologi RSUD Saiful Anwar Kota Malang.
2. Memeriksa keberadaan pemisahan sesuai bagian 2.1.4.2.
3. Menduga parameter menggunakan metode PMLE jika terdapat pemisahan, mengumpulkan data baru jika tidak terdapat pemisahan.
4. Menguji parameter secara simultan dan parsial menggunakan statistik uji pada persamaan 2.28 dan 2.31.
5. Memilih model terbaik menggunakan metode *Backward* sesuai bagian 2.1.7.
6. Menguji kelayakan model menggunakan statistik uji *Pearson* pada persamaan 2.32.
7. Menginterpretasi model regresi logistik multinomial menggunakan odds ratio sesuai bagian 2.1.8.

Diagram alir prosedur analisis ini disajikan pada Gambar 3.1.

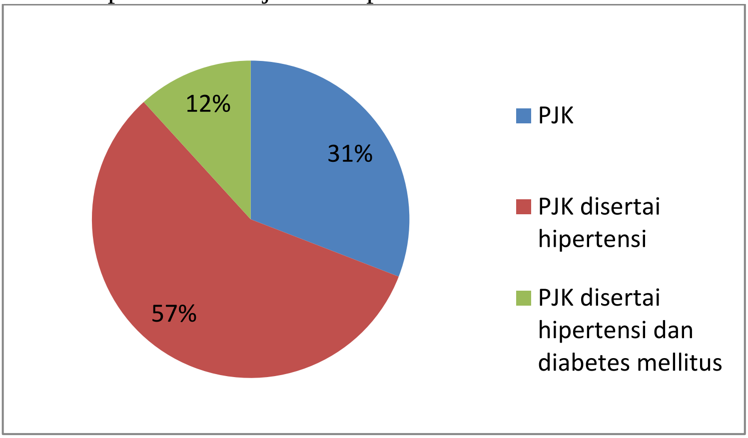


Gambar 3.1. Diagram Alir Prosedur Analisis Data

BAB IV
HASIL DAN PEMBAHASAN

4.1. Statistika Deskriptif

Gambaran umum data sebelum melakukan pemeriksaan keberadaan pemisahan dijelaskan pada Gambar 4.1.



Gambar 4.1. Diagram Kue Pie PJK

Penelitian ini dilakukan oleh Wibowo (2001) di poli kardiologi RSUD Saiful Anwar Kota Malang terhadap 68 pasien penderita Penyakit Jantung Koroner. Gambar 4.1 memperlihatkan bahwa persentase pasien penderita PJK disertai hipertensi (57%) lebih besar dibandingkan dengan penderita PJK (31%) dan PJK disertai hipertensi dan diabetes mellitus (12%).

4.2. Pemeriksaan Keberadaan Pemisahan

Berikut ini adalah hasil pendugaan parameter model profil pasien poli kardiologi RSUD Saiful Anwar Kota Malang dengan metode MLE.

Tabel 4.1. Penduga Parameter Model

t	Penduga Parameter								
	Intersep 1	Intersep 2	X_1	X_2	X_3	X_4	X_5	X_6	X_7
0	-0.806	2.015	0	0	0	0	0	0	0
1	4.987	7.808	0.011	-0.033	-0.019	-0.039	0.029	0.039	0.008
2	6.798	11.669	0.016	-0.047	-0.027	-0.069	0.053	0.069	0.015
3	8.808	15.539	0.016	-0.062	-0.029	-0.092	0.071	0.089	0.019
4	10.273	18.304	0.014	-0.074	-0.024	-0.104	0.081	0.099	0.022
5	10.901	19.524	0.013	-0.083	-0.018	-0.108	0.084	0.101	0.023
6	11.149	19.929	0.013	-0.086	-0.016	-0.108	0.084	0.1	0.023
7	11.268	20.098	0.014	-0.087	-0.015	-0.107	0.084	0.099	0.023
8	**11.321**	**20.173**	**0.015**	**-0.088**	**-0.015**	**-0.107**	**0.084**	**0.099**	**0.023**
9	11.343	20.204	0.015	-0.089	-0.015	-0.107	0.084	0.099	0.023
10	11.351	20.216	0.015	-0.089	-0.015	-0.107	0.084	0.099	0.023
11	11.355	20.221	0.015	-0.089	-0.015	-0.107	0.084	0.099	0.023
12	11.356	20.222	0.015	-0.089	-0.015	-0.107	0.084	0.099	0.023

Kriteria kekonvergenan = $\left|\widehat{\boldsymbol{\beta}}^{(t+1)} - \widehat{\boldsymbol{\beta}}^{(t)}\right| < 1 \times 10^{-6}$.

Hasil pendugaan parameter model dengan metode MLE pada Tabel 4.1 memperlihatkan pada iterasi ke delapan (t=8) kriteria kekonvergenan tidak terpenuhi, sehingga diperlukan pemeriksaan lanjutan terhadap peluang ketepatan alokasi setiap pengamatan untuk t>8.

Tabel 4.2. Peluang Ketepatan Alokasi Terbesar untuk t>8

Iterasi (t)	Peluang Ketepatan Alokasi	
	Terkecil	Terbesar
9	0.999252	1
10	0.999267	1
11	0.999976	1
12	0.124466	1

Mulai t=9 sampai t=12, nilai peluang ketepatan alokasi terbesar yang ditunjukkan dalam Tabel 4.2 lebih dari 0.95. Hal ini mengindikasikan bahwa tidak terjadi pemisahan sempurna melainkan pemisahan kurang sempurna karena nilai peluang ketepatan alokasi tidak sama dengan 1 untuk semua pasien. Untuk memastikan hal ini, dilakukan pemeriksaan ragam penduga parameter dibakukan yang ditampilkan pada Tabel 4.3.

Tabel 4.3. Ragam Penduga Parameter Dibakukan untuk t>8

Ragam Penduga	Iterasi (t)			
	9	10	11	12
$Var(\hat{\beta}_{Int\ 1})$	25.1369	25.1436	25.1462	25.1471
$Var(\hat{\beta}_{Int\ 2})$	36.0906	36.1145	36.1237	36.1273
$Var(\hat{\beta}_{Z_{X1}})$	0.0022	0.0022	0.0023	0.0022
$Var(\hat{\beta}_{Z_{X2}})$	0.0013	0.0013	0.0013	0.0013
$Var(\hat{\beta}_{Z_{X3}})$	0.0026	0.0027	0.0027	0.0027
$Var(\hat{\beta}_{Z_{X4}})$	0.0006	0.0006	0.0006	0.0006
$Var(\hat{\beta}_{Z_{X5}})$	0.0015	0.0016	0.0016	0.0016
$Var(\hat{\beta}_{Z_{X6}})$	0.0007	0.0007	0.0007	0.0007
$Var(\hat{\beta}_{Z_{X7}})$	0.0000	0.0000	0.0000	0.0000

Ragam penduga parameter dibakukan pada Tabel 4.3 memperlihatkan bahwa ragam penduga intersep 1 dan 2 lebih dari 5. Hal ini menguatkan keberadaan pemisahan kurang sempurna. Maka pendugaan parameter dilakukan dengan metode PMLE.

4.3. Pendugaan dan Pengujian Parameter Secara Simultan dan Parsial

Hasil pendugaan dan pengujian parameter secara simultan dan parsial ditampilkan pada Tabel 4.4.

Tabel 4.4. Hasil Pendugaan dan Pengujian Parameter

k	Peubah (X_j)	$\hat{\beta}_{kj}$	$SE(\hat{\beta}_{kj})$	W_{kj}	Nilai-p
2	X_0	-41.034	20.642	-1.987	0.046*
	X_1	-0.102	0.103	-0.995	0.319
	X_2	0.518	0.260	1.997	0.045*
	X_3	0.016	0.128	0.127	0.898
	X_4	0.123	0.079	1.556	0.119
	X_5	-0.107	0.118	-0.906	0.364
	X_6	-0.111	0.059	-1.877	0.060
	X_7	-0.067	0.030	-2.211	0.027*
3	X_0	-113.359	56.099	-2.020	0.043*
	X_1	-1.021	0.629	-1.622	0.104
	X_2	1.784	0.874	2.040	0.041*
	X_3	-2.153	1.340	-1.606	0.108
	X_4	1.092	0.587	1.859	0.062
	X_5	-0.315	0.213	-1.484	0.137
	X_6	-0.895	0.483	-1.851	0.064
	X_7	-0.245	0.123	-1.990	0.046*
G = 112.669; $P(\chi^2_{(7)} > G) < 0.0001$					

Hasil pendugaan dan pengujian parameter pada Tabel 4.4 memperlihatkan bahwa secara simultan terdapat paling sedikit satu peubah prediktor penentu tingkatan penyakit jantung koroner. Secara parsial, prediktor usia, tekanan darah *diastole*, kolesterol, HDL dan LDL bukan faktor penentu status penyakit pasien poli kardiologi untuk PJK disertai hipertensi dan PJK disertai hipertensi dan diabetes mellitus. Dengan demikian perlu dilakukan pemilihan model terbaik.

4.4. Pemilihan Model Terbaik

Pemilihan model terbaik menggunakan metode *backward* ditampilkan pada Tabel 4.5.

Tabel 4.5. Pemilihan Model Terbaik

r	Peubah Prediktor yang Direduksi	G_{p-r+1}	G_{p-r}	G_r^2	$P(\chi^2_{(1)} > G_r^2)$	Keterangan
0	-	-	112.669	-	-	-
1	X_3	112.669	103.116	9.553	0.002*	Model tidak sesuai
1	X_5	112.669	112.526	0.143	0.705	Model sesuai
1	X_1	112.669	104.667	8.002	0.005*	Model tidak sesuai
1	X_4	112.669	83.292	29.377	< 0.001	Model tidak sesuai
1	X_6	112.669	94.438	18.231	< 0.001	Model tidak sesuai
2	X_5, X_3	112.526	97.530	14.996	< 0.001	Model tidak sesuai
2	X_5, X_1	112.526	101.940	10.586	< 0.001	Model tidak sesuai
2	X_5, X_4	112.526	80.24	32.286	< 0.001	Model tidak sesuai
2	X_5, X_6	112.526	94.37	18.156	< 0.001	Model tidak sesuai

Pemilihan model terbaik pada Tabel 4.5 menunjukkan bahwa ketika prediktor X_3, X_1, X_4 dan X_6 dihilangkan dari model akan mengakibatkan model tidak sesuai. Sedangkan jika prediktor HDL(X_5) dihilangkan, model akan sesuai. Dengan demikian, model terbaik mengandung prediktor usia, tekanan darah *sistole*, tekanan darah *diastole*, kolesterol, LDL dan Trigliserid.

Tabel 4.6 menyajikan hasil pendugaan parameter model terbaik.

Tabel 4.6. Hasil Pendugaan Parameter Model Terbaik

Peubah (X_j)	$\hat{\beta}_{kj}$	
	$\hat{\beta}_{2j}$	$\hat{\beta}_{3j}$
X_0 (Intersep)	-35.206	-172.863
X_1 (Usia)	-0.104	-2.500
X_2 (Tekanan darah *sistole*)	0.326	3.465
X_3 (Tekanan darah *diastole*)	0.069	-4.869
X_4 (Kolesterol)	0.080	1.959
X_6 (LDL)	-0.098	-1.506
X_7 (Trigliserid)	-0.034	-0.501

Berdasarkan hasil pendugaan parameter model terbaik pada Tabel 4.6 didapatkan dua model regresi logistik multinomial terbaik:

a. Logit ke-1:

$$\hat{g}_2(\boldsymbol{x}) = -35.206 - 0.104x_1 + 0.326x_2 + 0.069x_3 + 0.08x_4 - 0.098x_6 - 0.034x_7$$

b. Logit ke-2:

$$\hat{g}_3(\boldsymbol{x}) = -172.863 - 2.5x_1 + 3.465x_2 - 4.869x_3 + 1.959x_4 - 1.506x_6 - 0.501x_7$$

dan peluang model regresi logistik multinomial terbaik untuk:

a. Penyakit jantung koroner:

$$\hat{\pi}_1(\boldsymbol{x}) = \frac{1}{1+\exp(-35.206-0.104x_1+0.326x_2+0.069x_3+0.08x_4-0.098x_6-0.034x_7)+\exp(-172.863-2.5x_1+3.465x_2-4.869x_3+1.959x_4-1.506x_6-0.501x_7)}$$

b. Penyakit jantung koroner disertai hipertensi:

$$\hat{\pi}_2(\boldsymbol{x}) = \frac{\exp(-35.206-0.104x_1+0.326x_2+0.069x_3+0.08x_4-0.098x_6-0.034x_7)}{1+\exp(-35.206-0.104x_1+0.326x_2+0.069x_3+0.08x_4-0.098x_6-0.034x_7)+\exp(-172.863-2.5x_1+3.465x_2-4.869x_3+1.959x_4-1.506x_6-0.501x_7)}$$

c. Penyakit jantung koroner disertai hipertensi dan diabetes mellitus:

$$\hat{\pi}_3(\boldsymbol{x}) = \frac{\exp(-172.863-2.5x_1+3.465x_2-4.869x_3+1.959x_4-1.506x_6-0.501x_7)}{1+\exp(-35.206-0.104x_1+0.326x_2+0.069x_3+0.08x_4-0.098x_6-0.034x_7)+\exp(-172.863-2.5x_1+3.465x_2-4.869x_3+1.959x_4-1.506x_6-0.501x_7)}$$

4.5. Pengujian Kelayakan model

Hasil pengujian kelayakan model terbaik ditampilkan pada Tabel 4.7.

Tabel 4.7. Hasil Pengujian Kelayakan Model Terbaik

Statistik Uji $\chi^2_{Pearson}$	Nilai-p
12.686	1.000

Berdasarkan hasil pengujian kelayakan model terbaik pada Tabel 4.7 menunjukkan nila-p sebesar 1, berarti dengan taraf nyata 0.05 dapat disimpulkan bahwa model terbaik bagi data profil pasien poli kardiologi RSUD Saiful Anwar Kota Malang bersifat layak.

4.6. Interpretasi Model

Interpretasi penduga parameter model terbaik logit 1 dan 2 dilakukan terhadap odds ratio ditampilkan pada Tabel 4.8.

Tabel 4.8. Odds Ratio Penduga Parameter Model Terbaik

Peubah (X_j)	$\hat{\beta}_{kj}$			
	k=2		k=3	
	$\hat{\beta}_{2j}$	Odds Ratio	$\hat{\beta}_{3j}$	Odds Ratio
X_0 (Intersep)	-0.104	0.901	-2.500	0.082
X_1 (Usia)	0.326	1.385	3.465	31.976
X_2 (Tekanan darah *sistole*)	0.069	1.071	-4.869	0.007
X_3 (Tekanan darah *diastole*)	0.080	1.083	1.959	7.092
X_4 (Kolesterol)	-0.098	0.906	-1.506	0.221
X_6 (LDL)	-0.034	0.966	-0.501	0.605
X_7 (Trigliserid)	-0.104	0.901	-2.500	0.082

Odds ratio penduga parameter model terbaik pada Tabel 4.8 menghasilkan kesimpulan untuk logit 1:

- Setiap 1000 pasien penderita penyakit jantung koroner dengan pertambahan usia 1 tahun maka terdapat 901 pasien akan menderita penyakit jantung koroner disertai hipertensi.

- Setiap 1000 pasien penderita penyakit jantung koroner dengan kenaikan 10 mmHg tekanan darah *sistole* maka terdapat 1385 pasien akan menderita penyakit jantung koroner disertai hipertensi.
- Setiap 1000 pasien penderita penyakit jantung koroner dengan kenaikan 10 mmHg tekanan darah *diastole* maka terdapat 1071 pasien akan menderita penyakit jantung koroner disertai hipertensi.
- Setiap 1000 pasien penderita penyakit jantung koroner dengan kenaikan 1 mg/dl kolesterol maka terdapat 1083 pasien akan menderita penyakit jantung koroner disertai hipertensi.
- Setiap 1000 pasien penderita penyakit jantung koroner dengan kenaikan 1 mg/dl LDL maka terdapat 906 pasien akan menderita penyakit jantung koroner disertai hipertensi.
- Setiap 1000 pasien penderita penyakit jantung koroner dengan kenaikan 1 mg/dl trigliserid maka terdapat 966 pasien akan menderita penyakit jantung koroner disertai hipertensi.

dan untuk logit 2:

- Setiap 1000 pasien penderita penyakit jantung koroner dengan pertambahan usia 1 tahun maka terdapat 82 pasien akan menderita penyakit jantung koroner disertai hipertensi dan diabetes mellitus.
- Setiap 1000 pasien penderita penyakit jantung koroner dengan kenaikan 10 mmHg tekanan darah *sistole* maka terdapat 31976 pasien akan menderita penyakit jantung koroner disertai hipertensi dan diabetes mellitus.
- Setiap 1000 pasien penderita penyakit jantung koroner dengan kenaikan 10 mmHg tekanan darah *diastole* maka terdapat 7 pasien akan menderita penyakit jantung koroner disertai hipertensi dan diabetes mellitus.
- Setiap 1000 pasien penderita penyakit jantung koroner dengan kenaikan 1 mg/dl kolesterol maka terdapat 7092 pasien akan menderita penyakit jantung koroner disertai hipertensi dan diabetes mellitus.
- Setiap 1000 pasien penderita penyakit jantung koroner dengan kenaikan 1 mg/dl LDL maka terdapat 221 pasien akan menderita penyakit jantung koroner disertai hipertensi dan diabetes mellitus.

- Setiap 1000 pasien penderita penyakit jantung koroner dengan kenaikan 1 mg/dl trigliserid maka terdapat 605 pasien akan menderita penyakit jantung koroner disertai hipertensi dan diabetes mellitus.

BAB V
KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan beberapa hal berikut:

1) Model Regresi Logistik Multinomial terbaik yang terbentuk untuk:

   logit ke-1:

   $$\hat{g}_2(\boldsymbol{x}) = -35.206 - 0.104x_1 + 0.326x_2 + 0.069x_3 + 0.08x_4 - 0.098x_6 - 0.034x_7$$

   logit ke-2:

   $$\hat{g}_3(\boldsymbol{x}) = -172.863 - 2.5x_1 + 3.465x_2 - 4.869x_3 + 1.959x_4 - 1.506x_6 - 0.501x_7$$

   Peluang model regresi logistik multinomial terbaik untuk kategori:

   Penyakit jantung koroner:

   $$\hat{\pi}_1(\boldsymbol{x}) = \frac{1}{1+\exp(-35.206-0.104x_1+0.326x_2+0.069x_3+0.08x_4-0.098x_6-0.034x_7)+\exp(-172.863-2.5x_1+3.465x_2-4.869x_3+1.959x_4-1.506x_6-0.501x_7)}$$

   Penyakit jantung koroner disertai hipertensi:

   $$\hat{\pi}_2(\boldsymbol{x}) = \frac{\exp(-35.206-0.104x_1+0.326x_2+0.069x_3+0.08x_4-0.098x_6-0.034x_7)}{1+\exp(-35.206-0.104x_1+0.326x_2+0.069x_3+0.08x_4-0.098x_6-0.034x_7)+\exp(-172.863-2.5x_1+3.465x_2-4.869x_3+1.959x_4-1.506x_6-0.501x_7)}$$

   Penyakit jantung koroner disertai hipertensi dan diabetes mellitus:

   $$\hat{\pi}_3(\boldsymbol{x}) = \frac{\exp(-172.863-2.5x_1+3.465x_2-4.869x_3+1.959x_4-1.506x_6-0.501x_7)}{1+\exp(-35.206-0.104x_1+0.326x_2+0.069x_3+0.08x_4-0.098x_6-0.034x_7)+\exp(-172.863-2.5x_1+3.465x_2-4.869x_3+1.959x_4-1.506x_6-0.501x_7)}$$

2) Model terbaik bagi data profil pasien kardiologi RSSA Malang mengandung prediktor usia, tekanan darah *sistole*, tekanan darah *diastole*, kolesterol, LDL dan Trigliserid.
3) Data profil pasien poli kardiologi RSUD Saiful Anwar mengandung pemisahan kurang sempurna sehingga menggunakan metode PMLE.

5.2. Saran

Saran-saran yang dapat diberikan:

1) Agar peneliti lain menerapkan metode PMLE pada model regresi logistik ordinal.
2) Gunakan metode *Double Penalized Maximum Likelihood Estimation* (DPMLE) pada model regresi logistik multinomial jika data mengandung pemisahan dan multikolinieritas.

DAFTAR PUSTAKA

Adhita, P.M. 2010. *Perbedaan Angka Kejadian Hipertensi Antara Pria dan Wanita Penderita DM Berusia Lebih Dari 45 Tahun*. Skripsi Fakultas Kedokteran Universitas Muhammadiyah Surakarta. Surakarta.

Agresti, A. 2002. *Categorical Data Analysis*. John Wiley & Sons. New York.

Arieska, P.K. 2010. *Perbandingan Metode Maximum Likelihood Estimation Dengan Bayesian Pada Regresi Logistik Multinomial*. Skripsi. Institut Teknologi Sepuluh Nopember. Surabaya.

Albert, A. dan Anderson, J.A. 1984. *On the Existence of Maximum Likelihood Estimates in Logistic Regression Model*. Biometrika. Vol. 71: 1-10.

Allison, P. 2008. *Convergence Failures in Logistic Regression*. http://www2.sasa.com/proceedings/forum2008/360-2008.pdf diakses pada tanggal 7 Januari 2016.

Djohan, T.A.B. 2004. *Penyakit Jantung Koroner Dan Hipertensi*. http://library.usu.ac.id/download/fk/gizi-bahri10.pdf, diakses pada tanggal 21 Februari 2016.

Firth, D. 1993. *Bias Reduction of Maximum Likelihood Estimates*. Biometrika 80: 27-38.

Hariadi dan Ali, A.R. 2005. *Hubungan Obesitas dengan Beberapa Faktor Risiko PJK Di Laboratorium Klinik Prodia Makasar*. http://arali2008.files.wordpress.com/2008/09/obesitas-dan-jantung-koroner.pdf, diakses pada tanggal 21 Februari 2016.

Heinze, G. dan Schemper, M. 2002. *A Solution to The Problem of Separation in Logistic Regression*. Statistics in Medicine 21: 2409-2419.

Hewson, P. 2009. *The Jeffreys Prior*. http://users.aims.ac.za/~paulhewson/jeffreysprior.pdf diakses pada tanggal 7 Januari 2016.

Hosmer, D.W. dan Lemeshow, S. 2000. *Applied Logistic Regression*. John Wiley & Sons. New York.

Lusiana, E.D. 2012. *Penerapan Metode Penalized Maximum Likelihood Estimation Untuk Mengatasi Pemisahan (Separation) Pada Analisis Regresi*. Skripsi. Universitas Brawijaya. Malang.

Majid, C. 2010. *Trigliserida*. http://www.triglislerida.com, diakses pada tanggal 21 Februari 2016.

McCullagh, P. dan Nelder, J.A. 1989. *Generalized Linear Models*. Chapman and Hall. London.

Walpole, R.E. 1995. *Pengantar Statistika*. PT Gramedia Pustaka Utama. Jakarta.

Wibowo, A.P. 2001. *Hubungan Episode Depresi Dan Kecemasan Dengan Faktor-Faktor Risiko Terutama Infeksi Heliobacter Pylori Pada Penyakit Jantung Koroner Di Poli Cardiologi RSSA Malang.* Skripsi Jurusan Kedokteran Fakultas Kedokteran Universitas Brawijaya, tidak dipublikasikan.

LAMPIRAN

Lampiran 1. Data Profil Pasien Poli Kardiologi RSUD Saiful Anwar Kota Malang

No	Y	X1	X2	X3	X4	X5	X6	X7
1	1	68	100	70	187	19.5	115.9	388
2	1	69	100	70	179	30.9	120.9	266
3	1	71	110	80	184	55.9	135.9	91
4	1	69	100	60	235	61.8	50	246
5	1	62	110	70	187	38.3	129.7	225
6	1	47	105	65	183	38.5	123.3	236
7	1	62	100	70	128	44.8	79.6	148
8	1	61	110	80	196	35.4	51.4	676
9	1	49	110	80	183	32.4	135.2	207
10	1	70	100	70	104	52	39.8	141
11	1	52	110	80	211	52.6	138.2	231
12	1	53	100	60	234	44.6	178.4	185
13	1	65	100	70	123	24	100.8	121
14	1	65	120	80	146	50.7	104.7	83
15	1	38	110	80	177	44.7	126.5	159
16	1	54	110	80	289	37	135.8	711
17	1	57	110	70	180	53.9	132.3	99
18	1	51	120	80	256	61.5	170.1	252
19	1	51	110	50	139	44.8	101	96
20	1	58	110	90	202	45.5	153.5	145
21	1	49	110	80	215	44.7	142.7	268
22	2	61	140	80	271	46	157.8	336
23	2	50	150	100	191	31.2	111.2	243
24	2	56	160	100	236	31.2	143.8	305
25	2	68	140	90	144	40.3	60.1	218
26	2	69	140	90	250	93.7	103.5	264
27	2	44	110	80	220	53.3	140.3	132
28	2	58	120	100	230	47.4	142.4	201
29	2	57	160	110	226	28.6	167.2	151
30	2	75	200	100	232	65.6	121.4	225
31	2	79	120	80	178	58.3	96.1	118
32	2	49	140	90	202	59.1	113.5	147
33	2	65	130	80	162	27.9	91.1	125

Lampiran 1 (Lanjutan)

No	Y	X1	X2	X3	X4	X5	X6	X7
34	2	54	130	100	240	28.6	172.2	196
35	2	60	140	90	209	34.4	133.2	207
36	2	65	150	100	238	29.9	154.7	267
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
52	2	58	140	80	198	49	105.6	217
53	2	58	130	90	143	42.9	67.1	165
54	2	63	130	100	177	24	133.2	99
55	2	67	180	100	223	40.8	129.4	264
56	2	61	170	100	245	32	140	44.9
57	2	41	130	90	268	50.7	148.3	345
58	2	50	150	100	241	37.7	166.9	182
59	2	54	130	90	180	35.4	122.2	112
60	2	65	160	100	254	39	172.2	214
61	3	51	150	90	287	48.1	148.3	203
62	3	66	170	90	332	48.7	176.1	286
63	3	74	180	100	263	61.7	114.1	186
64	3	74	150	70	244	51.7	94.1	241
65	3	58	160	100	337	59.1	182.5	227
66	3	65	150	80	197	54.6	72.6	99
67	3	44	160	90	291	45.5	154.3	206
68	3	73	150	80	297	39	180.4	138

Keterangan:

X_1 : Usia (tahun)
X_2 : Tekanan darah *sistole* (mmHg)
X_3 : Tekanan darah *diastole* (mmHg)
X_4 : Kolesterol (mg/dl)
X_5 : HDL (mg/dl)
X_6 : LDL (mg/dl)
X_7 : Trigliserid (mg/dl)

Lampiran 2. Skor baku Profil Pasien Poli Kardiologi RSUD Saiful Anwar Kota Malang

No	Y	Z_{X1}	Z_{X2}	Z_{X3}	Z_{X4}	Z_{X5}	Z_{X6}	Z_{X7}
1	1	0.917419	-1.40185	-1.10034	-0.53029	-1.55962	-0.2053	1.330139
2	1	1.022266	-1.40185	-1.10034	-0.67709	-0.93074	-0.06809	0.362265
3	1	1.231962	-0.97148	-0.38089	-0.58534	0.448375	0.343574	-1.02608
4	1	1.022266	-1.40185	-1.81979	0.350555	0.773847	-2.01386	0.203597
5	1	0.288332	-0.97148	-1.10034	-0.53029	-0.52252	0.173421	0.036995
6	1	-1.28439	-1.18666	-1.46006	-0.60369	-0.51149	-0.00222	0.124263
7	1	0.288332	-1.40185	-1.10034	-1.61299	-0.16395	-1.20152	-0.57388
8	1	0.183484	-0.97148	-0.38089	-0.36513	-0.6825	-1.97544	3.614957
9	1	-1.07469	-0.97148	-0.38089	-0.60369	-0.848	0.324363	-0.10581
10	1	1.127114	-1.40185	-1.10034	-2.05341	0.233233	-2.29379	-0.62941
11	1	-0.76015	-0.97148	-0.38089	-0.08987	0.266332	0.406695	0.084596
12	1	-0.6553	-1.40185	-1.81979	0.332205	-0.17499	1.509941	-0.28034
13	1	0.602875	-1.40185	-1.10034	-1.70474	-1.31138	-0.61971	-0.78808
14	1	0.602875	-0.54112	-0.38089	-1.28267	0.161519	-0.51268	-1.08955
15	1	-2.22802	-0.97148	-0.38089	-0.71379	-0.16947	0.085601	-0.48661
16	1	-0.55045	-0.97148	-0.38089	1.341502	-0.59424	0.340829	3.892626
17	1	-0.23591	-0.97148	-1.10034	-0.65874	0.338046	0.244776	-0.96261
18	1	-0.86499	-0.54112	-0.38089	0.735923	0.757298	1.282156	0.251197
19	1	-0.86499	-0.97148	-2.53924	-1.41113	-0.16395	-0.61422	-0.98641
20	1	-0.13106	-0.97148	0.338565	-0.25502	-0.12534	0.826587	-0.59768
21	1	-1.07469	-0.97148	-0.38089	-0.01646	-0.16947	0.530192	0.378131
22	2	0.183484	0.319608	-0.38089	1.011186	-0.09776	0.944596	0.917602
23	2	-0.96984	0.749972	1.058016	-0.45688	-0.91419	-0.33429	0.179796
24	2	-0.34076	1.180335	1.058016	0.368906	-0.91419	0.560381	0.671667
25	2	0.917419	0.319608	0.338565	-1.31937	-0.41219	-1.73668	-0.01854

Lampiran 2 (Lanjutan)

26	2	1.022266	0.319608	0.338565	0.625818	2.533601	-0.54561	0.346398
27	2	-1.59893	-0.97148	-0.38089	0.075292	0.304947	0.464327	-0.70081
28	2	-0.13106	-0.54112	1.058016	0.258801	-0.02052	0.521959	-0.15341
29	2	-0.23591	1.180335	1.777467	0.185398	-1.05762	1.202569	-0.55008
30	2	1.651353	2.901789	1.058016	0.295503	0.983473	-0.05436	0.036995
31	2	2.070745	-0.54112	-0.38089	-0.69544	0.580771	-0.74869	-0.81188
32	2	-1.07469	0.319608	0.338565	-0.25502	0.624903	-0.27117	-0.58181
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
60	2	0.707723	1.610698	0.338565	2.130589	0.05119	1.44682	0.520933
61	3	1.546506	2.041062	1.058016	0.864379	0.768331	-0.2547	-0.27241
62	3	1.546506	0.749972	-1.10034	0.515713	0.216684	-0.80358	0.16393
63	3	-0.13106	1.180335	1.058016	2.222343	0.624903	1.622461	0.052862
64	3	0.602875	0.749972	-0.38089	-0.34678	0.376661	-1.39363	-0.96261
65	3	-1.59893	1.180335	0.338565	1.378203	-0.12534	0.848542	-0.11374
66	3	1.441658	0.749972	-0.38089	1.488309	-0.48391	1.564828	-0.65321
67	3	0.917419	-1.40185	-1.10034	-0.53029	-1.55962	-0.2053	1.330139
68	3	1.022266	-1.40185	-1.10034	-0.67709	-0.93074	-0.06809	0.362265
$\bar{Z}_{X_j}$		0	0	0	0	0	0	0
S_Z^2		1	1	1	1	1	1	1

Lampiran 3. Struktur Data Peubah Respon Tiga Kategori

i	$Y_{i,k}$		
	k		
	1	2	3
1	1	0	0
2	1	0	0
3	1	0	0
. . .	. . .	. . .	. . .
n_1	1	0	0
1	0	1	0
2	0	1	0
3	0	1	0
. . .	. . .	. . .	. . .
n_2	0	1	0
1	0	0	1
2	0	0	1
3	0	0	1
. . .	. . .	. . .	. . .
n_3	0	0	1

Keterangan:

n_1 : Banyaknya respon berkategori 1
n_2 : Banyaknya respon berkategori 2
n_3 : Banyaknya respon berkategori 3
di mana $\sum_{k=1}^{q=3} n_k = n$
n : Banyaknya semua pengamatan

Lampiran 4. Luaran Statistika Deskriptif

Summary statistics:

Variable	Categories	Frequencies	%
Y	1	21	30.882
	2	39	57.353
	3	8	11.765

Variable	Observations	Minimum	Maximum	Mean	Std. deviation
X1	68	36.000	79.000	59.250	9.538
X2	68	100.000	200.000	132.574	23.236
X3	68	50.000	140.000	85.294	13.899
X4	68	104.000	381.000	215.897	54.493
X5	68	19.500	128.000	47.772	18.128
X6	68	39.800	182.500	123.381	36.438
X7	68	44.900	711.000	220.337	126.049

Lampiran 5. Luaran Pendugaan Parameter dengan Metode MLE

t	-2 Log L	intercept 0	intercept 1	x1	x2	x3	x4	x5	x6	x7
0	126.954	-0.806	2.015	0	0	0	0	0	0	0
1	68.825	4.987	7.808	0.011	-0.033	-0.019	-0.039	0.029	0.039	0.008
2	46.035	6.798	11.669	0.016	-0.047	-0.027	-0.069	0.053	0.069	0.015
3	39.905	8.808	15.539	0.016	-0.062	-0.029	-0.092	0.071	0.089	0.019
4	38.639	10.273	18.304	0.014	-0.074	-0.024	-0.104	0.081	0.099	0.022
5	38.453	10.901	19.524	0.013	-0.083	-0.018	-0.108	0.084	0.101	0.023
6	38.421	11.149	19.929	0.013	-0.086	-0.016	-0.108	0.084	0.100	0.023
7	38.414	11.268	20.098	0.014	-0.088	-0.015	-0.107	0.084	0.099	0.023
8	38.413	11.321	20.173	0.014	-0.088	-0.015	-0.107	0.084	0.099	0.023
9	38.412	11.343	20.204	0.015	-0.089	-0.015	-0.107	0.084	0.099	0.023
10	38.412	11.351	20.216	0.015	-0.089	-0.015	-0.107	0.084	0.099	0.023
11	38.412	11.355	20.221	0.015	-0.089	-0.015	-0.107	0.084	0.099	0.023
12	38.412	11.356	20.222	0.015	-0.089	-0.015	-0.107	0.084	0.099	0.023

Lampiran 6. Peluang Ketepatan Alokasi Terbesar

Iterasi	-2 Log L	Peluang Ketepatan Alokasi Terbesar
9	38.412	1
10	38.412	1
11	38.412	1
12	38.412	1

Lampiran 7. Matriks Ragam-peragam Penduga Parameter Peubah Prediktor Dibakukan t>8

t = 9

	$I1$	$I2$	$ZX1$	$ZX2$	$ZX3$	$ZX4$	$ZX5$	$ZX6$	$ZX7$
Intercept 1	25.136	28.586	−0.141	−0.022	−0.080	−0.023	−0.039	−0.007	0.001
Intercept 2	28.586	36.090	−0.145	−0.067	−0.049	−0.058	−0.007	0.020	0.007
$ZX1$	−0.141	−0.145	0.002	−0.001	0.000	0.000	−0.000	0.000	0.000
$ZX2$	−0.022	−0.067	−0.001	0.001	−0.001	0.000	−0.000	−0.000	−0.000
$ZX3$	−0.080	−0.049	0.001	−0.001	0.002	0.000	0.000	−0.000	0.000
$ZX4$	−0.023	−0.058	0.000	0.000	0.000	0.001	−0.001	−0.001	−0.000
$ZX5$	−0.039	−0.007	−0.000	−0.000	0.000	−0.001	0.002	0.001	0.000
$ZX6$	−0.007	0.020	0.000	−0.000	−0.000	−0.001	0.001	0.001	0.000
$ZX7$	0.001	0.007	0.000	−0.000	0.000	−0.000	0.000	0.000	0.000

t = 10

	$I1$	$I2$	$ZX1$	$ZX2$	$ZX3$	$ZX4$	$ZX5$	$ZX6$	$ZX7$
Intercept 1	25.143	28.591	−0.141	−0.022	−0.080	−0.023	−0.039	−0.007	0.001
Intercept 2	28.591	36.114	−0.145	−0.067	−0.049	−0.058	−0.007	0.020	0.007
$ZX1$	−0.141	−0.145	0.002	−0.001	0.000	0.000	−0.000	0.000	0.000
$ZX2$	−0.022	−0.067	−0.001	0.001	−0.001	0.000	−0.000	−0.000	−0.000
$ZX3$	−0.080	−0.049	0.000	−0.001	0.002	0.000	0.000	−0.000	0.000
$ZX4$	−0.023	−0.058	0.000	0.000	0.000	0.001	−0.001	−0.001	−0.000
$ZX5$	−0.039	−0.007	−0.000	−0.000	0.000	−0.001	0.002	0.001	0.000
$ZX6$	−0.007	0.020	0.000	−0.000	−0.000	−0.001	0.001	0.001	0.000
$ZX7$	0.001	0.007	0.000	−0.000	0.000	−0.000	0.000	0.000	0.000

t = 11

	$I1$	$I2$	$ZX1$	$ZX2$	$ZX3$	$ZX4$	$ZX5$	$ZX6$	$ZX7$
Intercept 1	25.146	28.603	−0.141	−0.022	−0.080	−0.023	−0.039	−0.007	0.001
Intercept 2	28.603	36.123	−0.145	−0.067	−0.049	−0.058	−0.007	0.020	0.007
$ZX1$	−0.141	−0.145	0.002	−0.001	0.000	0.000	−0.000	0.000	0.000
$ZX2$	−0.022	−0.067	−0.001	0.001	−0.001	0.000	−0.000	−0.000	−0.000
$ZX3$	−0.080	−0.049	0.000	−0.001	0.002	0.000	0.000	−0.000	0.000
$ZX4$	−0.023	−0.058	0.000	0.000	0.000	0.001	−0.001	−0.001	−0.000
$ZX5$	−0.039	−0.007	−0.000	−0.000	0.000	−0.001	0.002	0.001	0.000
$ZX6$	−0.007	0.020	0.000	−0.000	−0.000	−0.001	0.001	0.001	0.000
$ZX7$	0.001	0.007	0.000	−0.000	0.000	−0.000	0.000	0.000	0.000

Lampiran 7 (Lanjutan)

t = 12

$$
\begin{array}{r c}
 & \begin{array}{ccccccccc} I1 & I2 & ZX1 & ZX2 & ZX3 & ZX4 & ZX5 & ZX6 & ZX7 \end{array} \\
\begin{array}{r} Intercept\ 1 \\ Intercept\ 2 \\ ZX1 \\ ZX2 \\ ZX3 \\ ZX4 \\ ZX5 \\ ZX6 \\ ZX7 \end{array} &
\begin{pmatrix}
25.147 & 28.605 & -0.141 & -0.022 & -0.080 & -0.023 & -0.039 & -0.007 & 0.001 \\
28.605 & 36.127 & -0.145 & -0.067 & -0.049 & -0.058 & -0.007 & 0.020 & 0.007 \\
-0.141 & -0.145 & 0.002 & -0.001 & 0.000 & 0.000 & -0.000 & 0.000 & 0.000 \\
-0.022 & -0.067 & -0.001 & 0.001 & -0.001 & 0.000 & -0.000 & -0.000 & -0.000 \\
-0.080 & -0.049 & 0.000 & -0.001 & 0.002 & 0.000 & 0.000 & -0.000 & 0.000 \\
-0.023 & -0.058 & 0.000 & 0.000 & 0.000 & 0.001 & -0.001 & -0.001 & -0.000 \\
-0.039 & -0.007 & -0.000 & -0.000 & 0.000 & -0.001 & 0.002 & 0.001 & 0.000 \\
-0.007 & 0.020 & 0.000 & -0.000 & -0.000 & -0.001 & 0.001 & 0.001 & 0.000 \\
0.001 & 0.007 & 0.000 & -0.000 & 0.000 & -0.000 & 0.000 & 0.000 & 0.000
\end{pmatrix}
\end{array}
$$

Lampiran 8. Luaran Pendugaan Parameter dengan Metode PMLE

Statistic	DF	Chi-square	Pr > Chi²
-2 Log(Likelihood)	14	112.670	< 0.0001
Score	14	75.492	< 0.0001
Wald	14	9.201	0.818

Model parameters (Variable Y):

Category	Source	Value	Standard error	Wald Chi-Square	Pr > Chi²
2.000	Intercept	-41.034	20.642	3.952	0.047
	X1	-0.102	0.103	0.991	0.319
	X2	0.518	0.260	2.600	0.107
	X3	0.016	0.128	0.016	0.898
	X4	0.123	0.079	2.423	0.120
	X5	-0.107	0.118	0.821	0.365
	X6	-0.111	0.059	3.526	0.060
	X7	-0.067	0.030	2.402	0.121
3.000	Intercept	-113.359	56.099	4.083	0.043
	X1	-1.021	0.629	2.632	0.105
	X2	1.784	0.874	4.164	0.041
	X3	-2.153	1.340	2.580	0.108
	X4	1.092	0.587	3.459	0.063
	X5	-0.315	0.213	2.203	0.138
	X6	-0.895	0.483	3.429	0.064
	X7	-0.245	0.123	3.961	0.047

Lampiran 9. Luaran Pemilihan Model Terbaik

r=1, tanpa X_3

Statistic	DF	Chi-square	Pr > Chi²
-2 Log(Likelihood)	12	103.116	< 0.0001
Score	12	68.107	< 0.0001
Wald	12	12.791	0.384

Model parameters (Variable Y):

Category	Source	Value	Standard error	Wald Chi-Square	Pr > Chi²
2.000	Intercept	-44.060	19.166	5.285	0.022
	X1	-0.110	0.103	1.130	0.288
	X2	0.459	0.207	4.912	0.027
	X4	0.121	0.072	2.779	0.096
	X5	-0.107	0.102	1.100	0.294
	X6	-0.105	0.060	3.053	0.081
	X7	-0.048	0.028	3.031	0.082
3.000	Intercept	-65.095	23.250	7.839	0.005
	X1	-0.102	0.133	0.583	0.445
	X2	0.524	0.212	6.126	0.013
	X4	0.256	0.092	7.733	0.005
	X5	-0.210	0.119	3.122	0.077
	X6	-0.228	0.083	7.507	0.006
	X7	-0.066	0.030	4.815	0.028

Lampiran 9 (Lanjutan)

r=1, tanpa X_5

Statistic	DF	Chi-square	Pr > Chi²
-2 Log(Likelihood)	12	112.526	< 0.0001
Score	12	65.707	< 0.0001
Wald	12	9.771	0.636

Model parameters (Variable Y):

Category	Source	Value	Standard error	Wald Chi-Square	Pr > Chi²
2.000	Intercept	-35.206	15.445	5.196	0.023
	X1	-0.104	0.094	1.229	0.268
	X2	0.326	0.176	3.422	0.064
	X3	0.069	0.115	0.361	0.548
	X4	0.080	0.041	3.736	0.053
	X6	-0.098	0.049	3.950	0.047
	X7	-0.034	0.019	3.190	0.074
3.000	Intercept	-172.863	89.386	3.740	0.053
	X1	-2.500	1.493	2.802	0.094
	X2	3.465	1.928	3.231	0.072
	X3	-4.869	3.016	2.607	0.106
	X4	1.959	1.138	2.963	0.085
	X6	-1.506	0.855	3.102	0.078
	X7	-0.501	0.280	3.192	0.074

Lampiran 9 (Lanjutan)

r=1, tanpa X_1

Statistic	DF	Chi-square	Pr > Chi²
-2 Log(Likelihood)	12	104.668	< 0.0001
Score	12	74.803	< 0.0001
Wald	12	10.051	0.611

Model parameters (Variable Y):

Category	Source	Value	Standard error	Wald Chi-Square	Pr > Chi²
2.000	Intercept	-40.729	17.973	5.135	0.023
	X2	0.331	0.182	3.313	0.069
	X4	0.099	0.062	2.553	0.110
	X6	-0.074	0.040	3.400	0.065
	X7	-0.035	0.023	2.441	0.118
	X5	-0.093	0.091	1.036	0.309
	X3	0.033	0.117	0.081	0.776
3.000	Intercept	-71.465	32.559	4.818	0.028
	X2	0.612	0.321	3.643	0.056
	X4	0.337	0.191	3.100	0.078
	X6	-0.250	0.150	2.766	0.096
	X7	-0.081	0.045	3.279	0.070
	X5	-0.202	0.133	2.315	0.128
	X3	-0.404	0.492	0.676	0.411

Lampiran 9 (Lanjutan)

r=1, tanpa X_4

Statistic	DF	Chi-square	Pr > Chi²
-2 Log(Likelihood)	12	83.292	< 0.0001
Score	12	54.586	< 0.0001
Wald	12	17.807	0.122

Model parameters (Variable Y):

Category	Source	Value	Standard error	Wald Chi-Square	Pr > Chi²
2.000	Intercept	-30.284	10.533	8.267	0.004
	X1	-0.066	0.074	0.802	0.370
	X2	0.243	0.098	6.154	0.013
	X6	-0.047	0.031	2.246	0.134
	X7	-0.007	0.007	1.005	0.316
	X5	0.075	0.075	0.991	0.319
	X3	0.115	0.086	1.790	0.181
3.000	Intercept	-35.373	11.182	10.008	0.002
	X1	-0.080	0.097	0.679	0.410
	X2	0.360	0.107	11.390	0.001
	X6	-0.004	0.038	0.014	0.907
	X7	-0.024	0.012	4.094	0.043
	X5	0.118	0.086	1.896	0.168
	X3	-0.083	0.094	0.780	0.377

Lampiran 9 (Lanjutan)

r=1, tanpa X_6

Statistic	DF	Chi-square	Pr > Chi²
-2 Log(Likelihood)	12	94.438	< 0.0001
Score	12	64.207	< 0.0001
Wald	12	17.298	0.139

Model parameters (Variable Y):

Category	Source	Value	Standard error	Wald Chi-Square	Pr > Chi²
2.000	Intercept	-33.636	11.202	9.017	0.003
	X1	0.010	0.065	0.022	0.883
	X2	0.254	0.102	6.210	0.013
	X4	0.025	0.031	0.647	0.421
	X7	-0.011	0.010	1.216	0.270
	X5	0.012	0.055	0.048	0.827
	X3	0.001	0.078	0.000	0.986
3.000	Intercept	-40.892	13.078	9.777	0.002
	X1	0.048	0.129	0.140	0.709
	X2	0.377	0.134	7.956	0.005
	X4	0.107	0.049	4.817	0.028
	X7	-0.051	0.022	5.187	0.023
	X5	0.004	0.082	0.002	0.963
	X3	-0.312	0.191	2.666	0.103

Lampiran 9 (Lanjutan)

r=2, tanpa X_5 dan X_3

Statistic	DF	Chi-square	Pr > Chi²
-2 Log(Likelihood)	10	97.53	< 0.0001

r=2, tanpa X_5 dan X_1

Statistic	DF	Chi-square	Pr > Chi²
-2 Log(Likelihood)	10	101.940	< 0.0001

r=2, tanpa X_5 dan X_4

Statistic	DF	Chi-square	Pr > Chi²
-2 Log(Likelihood)	10	80.24	< 0.0001

r=2, tanpa X_5 dan X_6

Statistic	DF	Chi-square	Pr > Chi²
-2 Log(Likelihood)	10	94.37	< 0.0001

Lampiran 10. Luaran Kelayakan Model

Goodness-of-Fit

	Chi-Square	Df	Sig.
Pearson	**12.686**	**122**	**1.000**

Lampiran 11. Luaran Odds Ratio

Category	Source	Value	Odds ratio	Odds ratio Lower bound (95%)	Odds ratio Upper bound (95%)
2.000	Intercept	-35.206			
	X1	-0.104	0.901	0.750	1.083
	X2	0.326	1.386	0.981	1.959
	X3	0.069	1.071	0.856	1.341
	X4	0.080	1.083	0.999	1.175
	X6	-0.098	0.907	0.823	0.999
	X7	-0.034	0.967	0.931	1.003
3.000	Intercept	-172.863			
	X1	-2.500	0.082	0.004	1.533
	X2	3.465	31.970	0.731	1397.958
	X3	-4.869	0.008	0.000	2.833
	X4	1.959	7.090	0.762	65.949
	X6	-1.506	0.222	0.042	1.185
	X7	-0.501	0.606	0.350	1.050